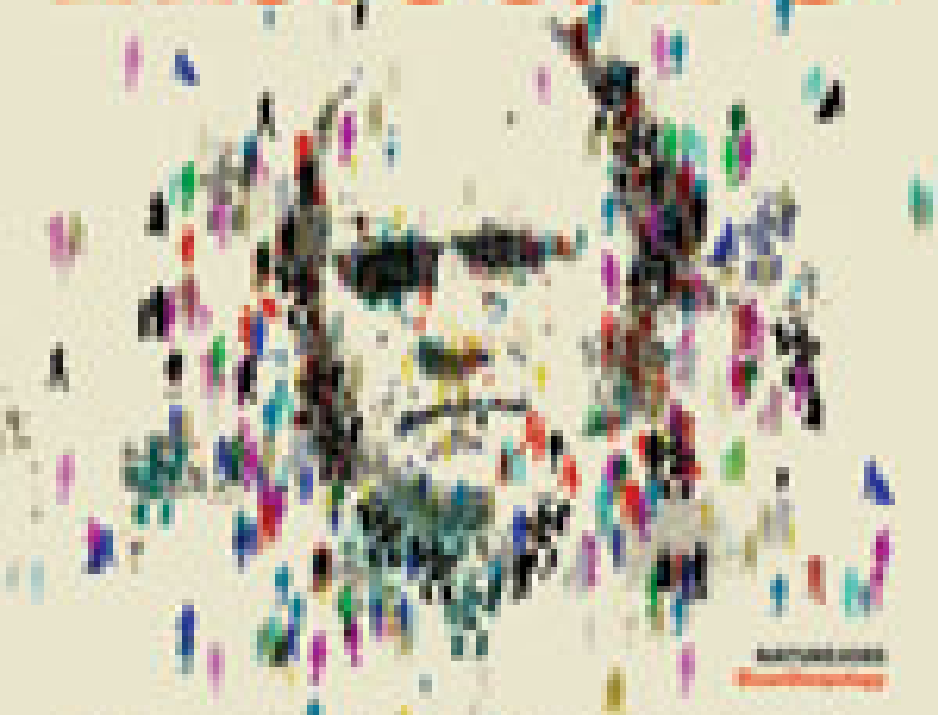
NATURE INSIGHT EVOLUTION

12 February 2009 www.nature.com/nature 279

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

nature



NATURE.COM/darwinday

EVERYBODY'S DARWIN

DARWIN 200 Humanity and evolution

- Seeing natural selection in the human genome
- Race and IQ: untouchable science?
- Pigeons and poetry

nature

www.nature.com/nature

Vol 457 | Issue no. 7231 | 12 February 2009

Humanity and evolution

Charles Darwin's thinking about the natural world was profoundly influenced by his revulsion for slavery.

Although history is not made entirely, or even mostly, by prominent men and women, two great exceptions to that rule were born exactly 200 years ago today, on 12 February 1809: Charles Darwin and Abraham Lincoln.

These men shared more than just a birthday, the loss of a mother in childhood and a date with immortality. They shared a position on one of the great issues of their age: the 'peculiar and powerful interest' of their fellow humans bound in slavery. When he circled the world in the 1830s, Darwin's delight at our planet's natural riches was repeatedly poisoned by the cruelties he saw meted out to slaves. "I thank God, I shall never again visit a slave-country," he wrote at the end of the *Voyage of the Beagle*.

A new historical study, *Darwin's Sacred Cause* by Adrian Desmond and James Moore (see page 792), seeks to unite Darwin's revulsion at slavery with his scientific work. It was common at the time to believe that the different races of men had been created separate and unequal. But the abolitionist beliefs that Darwin derived from his family, friends and social setting strongly disposed him to the idea that all men — Englishman and Hottentot, freeman and slave — were brothers united in shared ancestry. The ability to see that unity-in-variety was, Desmond and Moore argue, one of the things that allowed him to perceive something similar in the natural world as a whole. As Darwin wrote in an 1838 notebook, "I cannot help thinking good analogy might be traced between relationship of all men now living & the classification of animals." When Darwin sketched life's common descent as a family tree, it was because he believed in a family tree for humans — a belief in common kinship that was not a disinterested scientific finding, but rather an expression of moral and political persuasion. Darwin's thought always extended beyond the natural world. His ideas always had, and were meant to have, a social dimension.

Lessons from history

For all Darwin's noble ambitions, the century and a half since *On the Origin of Species* have shown how easily his image of a fiercely competitive world can be used to bolster pre-existing positions of power and privilege with buttresses of support that seem founded in an impartial consideration of the natural world. The history of arguments about humanity based on biology — both Darwin's biology and that of others who have come after — provides a sorry rehearsal of pretexts and apologias for everything from unthinking prejudice to forced sterilization and genocide (see page 786).

This history counsels caution as ever deeper and subtler forays into the science of human nature become possible. Deciphering the traces of natural selection in the human genome (see page 776), and dissecting the genetics of neurobiology and behaviour promise a new, more detailed and complex sense of how of how evolution has given human nature a definite biological form — while at the same time throwing new light on just how deeply biology can be influenced by society and culture. This is a rich field for research in both the natural and the social sciences, especially in the form of new collaborations between them (see page 780).

It is vital, however, that this new knowledge should be judged by far higher standards than the ideology passed off as biology that blighted so much of the twentieth century. Scientists have beliefs about what is right and wrong, just like everyone else. And try as they may to put them to one side — some try hard, some not so much — those beliefs will influence the way they do science, and the questions they ask and fail to ask. The scientific enterprise as a whole has to pay particular heed to the risk that preconceptions will creep in whenever what is being said about human nature has political or social implications.

This is particularly the case when science begins to look, as moral psychology is doing, at

Darwin200

For podcast and more online extras see www.nature.com/nature/darwin

the mechanisms by which people make decisions about right or wrong. Here it becomes peculiarly hard — and at the same time especially important — to resist the 'naturalistic fallacy' of inferring what ought to be from what is. Science may be able to tell us why some values are more easily held than others. But it cannot tell us whether taking the easy path in terms of which values we espouse is the right thing to do.

In fact, it provides us with a worked example to the contrary. The scientific endeavour itself is founded on values which natural selection would have seemed unlikely to foist on a bunch of violent, gregarious upright apes. Science tries to place no trust in authority; to some extent, society has to. Science tries to define its membership on the basis of inclusion, rather than exclusion; work on altruism suggests, worryingly, that communities more normally need an outgroup to form against. Science insists on the value of truth even when it is inconvenient or harmful; most people's beliefs tend to reinforce their self-interest.

In this unnaturalness lies the great strength of science. It is from this it derives its power as a way of understanding the world. And this is also what allows it, at its best, to resist, not reinforce, mores and prejudices that pose as truths of nature. This demanding, artificial code is what gives engaged, passionate and all-too-fallible human beings the collective power to produce results that are dispassionate, objective and reliable. And if science stays true to that code, it can act as a stern restraint on anyone seeking to go from the study of how people evolved to conclusions about how they should be treated now — to go, that is, against the values that both Darwin and Lincoln espoused.

Science can never prove humans alike in dignity, or equally deserving under the law; that is a truth that cannot be discovered. Like the ideals of malice towards none and charity towards all, it is something that must be made real through communal will. ■

Natural value

The economic downturn might be the best time to include ecosystem services in the real economy.

Worldwide momentum seems to be growing for an approach to environmental protection based on the 'ecosystem services' that nature provides for humans. These can range from watersheds filtering drinking water and forests sequestering carbon, to marshes dissipating the fury of storms. As long as the marketplace treats such services as free goods, proponents argue, the value of what nature does for humanity will effectively be set to zero and nature will continue to be trashed. But if the market could somehow be made to price the services appropriately, all those forests, streams, lakes, prairies and seashores would suddenly acquire real economic value, and people would have incentives to preserve them.

The ecosystem services approach clearly has great potential. Indeed, it is a natural extension of the market-based carbon tax or cap-and-trade approaches now being implemented to curb carbon emissions, in that it tackles environmental externalities historically ignored by the global economy. This month, moreover, a special issue of the Ecological Society of America's journal *Frontiers in Ecology and the Environment* offers some badly needed hard science on the subject. In one paper, for example, computer modelling of land use in the Willamette Basin in Oregon shows that commercial development ceases to be the most rational use of land when the simulation incorporates even conservatively priced payments for the carbon sequestration the land provides.

This special issue is just one of several recent developments that seem to herald ecosystem services' entry into mainstream scientific and political thinking. The Economics of Ecosystems and Biodiversity project funded by the European Union is collecting scientific evidence on the "economic consequences of biodiversity loss" until 31 March. The goal is to assemble a toolkit of techniques and information for those who want to do empirical ecosystem valuations of their own.

Elsewhere, last December the US Congress created an Office of Ecosystem Services and Markets in the Department of Agriculture, along with a government-wide Conservation and Land Management Environmental Services Board to advise on incentives for ecosystem services. And last November, a meeting in Malaysia kicked off a United Nations-sponsored Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services patterned after the agency's influential Intergovernmental Panel on Climate Change.

"Destroying ecosystems for short-term benefit is like killing the cow for its meat, when one might keep from starving by drinking its milk for years."

If such efforts are to succeed, however, and if ecosystem services are ever going to be successfully integrated into the regular economy, the scheme will have to be founded on even more hard science. Distributed sensing efforts such as the US National Science Foundation's National Ecological Observatory Network will be invaluable in this regard. But science policy-makers will need to make ecosystem monitoring, research, analysis and simulation a high priority in general — and on an ongoing basis. Granted, it will be difficult to find money for such activities in the current economic downturn. But they could provide a fair number of jobs. Monitoring tasks such as checking sediment traps and nitrogen levels in streams require many boots on the ground, for example, and streambed restoration requires many more.

The downturn also highlights perhaps the most worrying consequence of putting prices on the services provided by nature: it will make everything more expensive. This is not a politically palatable move at any time, much less now. But the whole point of the ecosystem services approach is that it saves everyone money in the long run. The damage that Hurricane Katrina inflicted on Louisiana in 2005 was a dramatic example of how ecosystem services — in this case, storm-buffering wetlands — are often cheaper to maintain than they are to live without. As governments launch large infrastructure projects to stimulate their economies, they should fold ecosystem services into the budget analyses. Destroying many ecosystems for short-term economic benefit is like killing the cow for its meat, when one might keep from starving by drinking its milk for years to come. Now is not the time to slaughter the cow. ■

RESEARCH HIGHLIGHTS

Relative values

Current Biol. doi:10.1016/j.cub.2008.12.049 (2009)

'Pygmy' is a blanket term for many Central African populations that have an average male height of 150 centimetres or less. Whether these groups share a common evolutionary divergence from other humans has been long-debated. But by comparing 28 genomic regions in 604 people — among them individuals from 9 Pygmy groups — Paul Verdu at the CNRS in Paris and his colleagues have confirmed the single-origin theory, and added to a part of recent human history.

The analysis suggests that Pygmies diverged from other human groups either 54,000 or 90,000 years ago (depending on whether the tallest Pygmy group, the Bongo, is included in statistical tests). They then remained as one interbreeding population until about 800 BC, when the agricultural expansion of their non-Pygmy neighbours seems to have driven a rapid genetic and cultural divergence among them.

For a longer story on this research, see http://tinyurl.com/awfk52

M. S. LEWIS/CORBIS

GEOSCIENCE

Ancient forests

J. Geophys. Res. doi:10.1029/2008JG000802 (2009)

Ancient polar forests may have influenced Earth's climate to a greater extent than previously thought, according to a study of five modern relatives of species that grew in them.

During four years of monitoring, David Beerling of the University of Sheffield, UK, Laura Llorens, now at the University of Girona in Spain, and their colleagues found that the trees emitted higher than expected levels of monoterpenes — compounds that produce secondary organic aerosols, which increase cloud cover. Emissions peaked during the prolonged periods of continuous light of the polar summer.

Given that the forests covered upwards of 40% of Earth's total land surface 145 million–65 million years ago, the group says that the monoterpenes the forests released could have had a large effect on the chemistry of the atmosphere and on climate.

GENETICS

Shady dealings

Science doi:10.1126/science.1165448 (2009)

The wolves prowling North America's forests may owe their success and the dark colour of their coats to the domesticated dog.

Greg Barsh of Stanford University in California and his colleagues looked at the DNA of 265 white, grey and black wolves and found that the mutation responsible for black coats seems to have been acquired from hybridization with man's best friend. They prefer this theory to that of dogs acquiring dark coats from wolves because the most recent common ancestor of wolves with the crucial K^B-bearing chromosomes lived later than the most recent common ancestor of dogs.

Because there is positive selection for K^B-bearing chromosomes in forest-dwelling wolves, domestic dogs may have contributed to the viability of wild wolf populations.

NANOTECHNOLOGY

Stretch marks

Nature Nanotechnol. doi:10.1038/nnano.2008.410 (2009)

Osteoarthritis is difficult to detect in its earliest stages because the joint cartilage shows no obvious signs of degeneration. Martin Stolz at Biozentrum, University of Basel, Switzerland, and his colleagues have now successfully applied atomic force microscopy to the problem, first prodding and then imaging cartilage from mouse and human joints.

SCIENCE/AAAS

This technique allowed them to pick up thickening and reductions in the elasticity of cartilage fibres long before any outward signs of osteoarthritis appeared. The team hopes that their research will lead to devices that detect the disease early and in a minimally invasive manner.

MOLECULAR BIOLOGY

Solo signal

Cell **136,** 411–419 (2009)

For skin, liver or stomach cells to behave like embryonic stem cells, researchers have so far had to add to them at least three genes that have been linked to cancer. But only one such gene, which encodes a molecule called Oct4 (and also Pou5F1), is needed to induce pluripotency — the ability to develop into most cell types — in mouse neural stem cells.

Unlike most successful attempts to reprogram cells, which take about three weeks, this approach by Hans Schöler of the Max Planck Institute for Molecular Biomedicine in Münster, Germany, and his co-workers took four to five weeks of culturing. Nonetheless, the efficiency of the reprogramming was roughly equal to that of mouse embryonic skin cells when Oct4 and three other reprogramming factors are added to them.

NEUROSCIENCE

Burning sensation

Nature Chem. Biol. doi:10.1038/nchembio.146 (2009)

Zinc activates an ion channel that also responds to pungent foods such as wasabi and cinnamon, possibly explaining why zinc overexposure causes pain and inflammation.

When ions flow through the TRPA1 channel, pain neurons fire. Ardem Patapoutian at the Genomics Institute of the Novartis Research Foundation in San Diego, California, and his colleagues report that 15% of mouse sensory neurons show increased activity when exposed to zinc, and that mouse neurons lacking TRPA1 are unaffected by the element. Mice without a working version of the gene that encodes TRPA1 displayed fewer signs of discomfort when injected with zinc acetate.

'Inside out' cell membranes were highly sensitive to zinc, suggesting that it activates TRPA1 from inside neurons.

ASTROPHYSICS

Where are the dwarfs?

Astron. J. **137,** 3009–3037 (2009)

A leading theory for the evolution of the Universe's structure explains how lumpy galaxies formed after matter was evenly distributed by a relatively smooth Big Bang. 'Cold dark matter' models predict that there should be many more small dwarf galaxies surrounding the Milky Way than have been identified. Kristin Chiboucas of the University of Hawaii in Honolulu and her colleagues think that the 'missing galaxy' problem is not just local; they report a paucity of dwarfs around Messier 81 (M81), a spiral galaxy in the constellation Ursa Major (the Great Bear).

A dozen of the 22 new dwarfs the team found lie within M81's gravitational influence. Theory predicted hundreds more. This finding adds weight to the evidence that missing dwarfs are a real problem and not one of imperfect observation.

CHEMISTRY

Bangs in the dark

Angew. Chem. Int. Edn doi:10.1002/anie.200804853 (2009)

Chemical hallmarks of certain explosives can be quickly identified by the loss of a characteristic glow in a very porous material.

The metal–organic framework (MOF) developed by Jing Li at Rutgers University in New Jersey and her colleagues is luminescent under ultraviolet light. But when its plentiful pores fill up with either 2,4-dinitrotoluene, a by-product of TNT manufacture, or 2,3-dimethyl-2,3-dinitrobutane (DMNB), a tracer molecule for plastic explosives, that luminescence fades.

This happens because the explosive molecules can quickly bind to the inside of the MOF, owing to its porosity, and once there they interfere with the electronic processes that otherwise cause luminescence.

ZOOLOGY

Horny?

Science **323,** 773–776 (2009)

Among many insects there are two male types, dominant and subordinate. But researchers studying the horns of dung beetles have found that some species have three types.

Mark Rowland at the University of New Mexico in Albuquerque and Douglas Emlen at the University of Montana in Missoula trapped and studied closely related dung beetles, the phanaeine. They noticed that males could have big horns, little horns or no horns at all. Some species had a developmental mechanism that entirely halted horn growth in subordinates, and others had one that merely attenuated horn growth. Five species had both — resulting in alpha, beta and gamma males (pictured below, clockwise from top right: female, gamma, beta and alpha).

SCIENCE/AAAS

NEUROSCIENCE

Precision memories

Nature Neurosci. doi:10.1038/nn.2263 (2009)

Why do some memories fade, whereas others retain their original clarity? Experiments on mice suggest that details of precise, long-term memories are stored outside the area of the brain involved in initial memory formation, the hippocampus.

Paul Frankland and his colleagues at the Hospital for Sick Children in Toronto, Canada, trained mice to expect an electric shock to their feet when placed in one type of chamber, but not when housed in a different-looking one. When returned to these chambers, the mice were more likely to freeze in the type associated with the foot shocks.

This behaviour persisted if the rodents' hippocampi were surgically damaged 42 days after training, but not if this was done just one day afterwards. So although longer-term memories may be stored outside the hippocampus, the storage process seems to be prolonged and hippocampus-dependent.

JOURNAL CLUB

Susan E. Hough
US Geological Survey, Pasadena, California

A seismologist considers a new method of earthquake prediction.

I am acutely aware that numerous methods of earthquake prediction at one time held great promise, but fell apart under proper scrutiny. In recent years, I have heard about many studies purporting to uncover evidence of electromagnetic precursors, almost all of which involved weak or non-existent statistical analysis.

But occasionally I come across research that is not so easy to dismiss. For example, data from the French micro-satellite DEMETER, which was launched in 2003 to investigate electromagnetic perturbations in the ionosphere, have been analysed by a team of French and Czech researchers (F. Němec *et al. Geophys. Res. Lett.* doi:10.1029/2007GRL032517; 2008). These authors find that there are very-low-frequency electromagnetic fluctuations in the ionosphere above the epicentres of moderate and large earthquakes that occur a day or two before the ground starts to shake.

Němec and colleagues' results could be fatally flawed. If electromagnetic disturbances are generated when earthquakes occur, what are apparently true signals of one earthquake could actually be signals related to a preceding shock. Or the analysis might go awry because of subtle data-selection biases. But if there are fatal flaws, they are not obvious.

In any case, as the authors themselves emphasize, the significance of the DEMETER results can be demonstrated only when data from many earthquakes are averaged. This highlights a key point: it is entirely possible for precursors to be real but of no use for prediction. If earthquake scientists can separate consideration of earthquake precursors from the highly charged debates about earthquake prediction, the research community might just learn something about earthquake processes.

Discuss this paper at **http://blogs.nature.com/nature/journalclub**

NEWS

Genome sequencing: the third generation

Companies unveil data from their latest technologies.

MARCO ISLAND, FLORIDA

Genome researchers gathered in a Florida hotel on 5 February, hoping to see whether companies that build 'third-generation' sequencing technologies can deliver on stunning claims such as sequencing human genomes in three minutes or selling them for $5,000. Although scientists were cautiously optimistic about the data unveiled, they still have major questions about how well this next generation of machines will work.

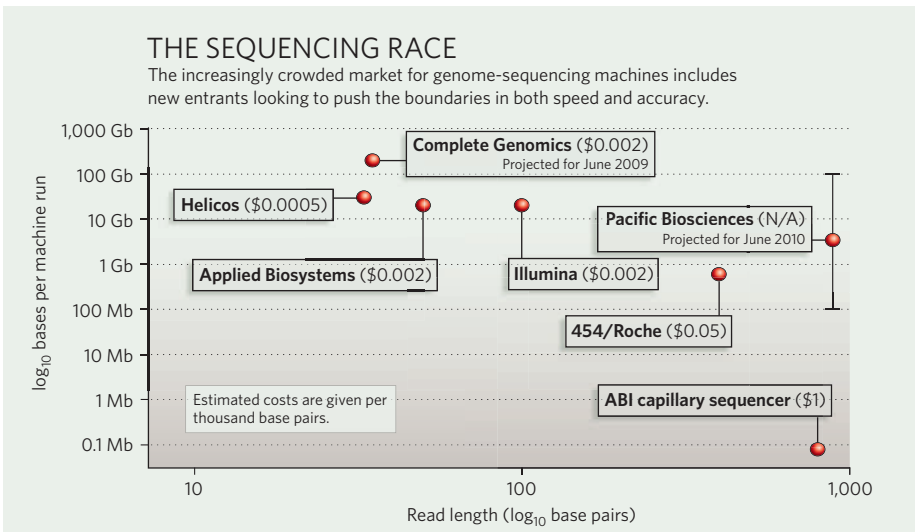
For one company, the Advances in Genome Biology and Technology meeting in Marco Island, Florida, was a major test. In October 2008, Complete Genomics of Mountain View, California, said it would sell whole human genomes in 2009 for $5,000, but it released no supporting data. At the conference, the company revealed a human genome it said it had sequenced using nine machines for eight days over Christmas.

The company's chief executive, Clifford Reid, says it assembled 254 gigabases (254 billion base pairs) of data into a draft covering 92% of the genome of an anonymous man, and that it read each base an average of 91 times. Like many of the high-speed sequencing technologies currently in use — usually called 'next-generation' technologies, as opposed to the third generation still in development — Complete Genomics produces short reads of DNA. By sequencing each base many times, it aims to diminish the potential errors that could creep in when the short reads are assembled into longer pieces. Reid says that the technology is highly accurate, with less than one-third of a per cent chance of making an error in any given base. That's comparable to the current generation of sequencers.

Complete Genomics is not selling sequencing machines, but instead performs all its work in-house on its own machines. This has made some scientists sceptical, but others have been encouraged by the company's data. "Their key thing is to show that they can have highly accurate base identification across the vast majority of the genome," says Chad Nusbaum, co-director of the sequencing centre at the Broad Institute in Cambridge, Massachusetts. "They're doing the right analyses, and it seems like things look pretty good."

"There's a lot of complexity in everybody's genome."

Speed and cost have been Complete Genomics' main selling points; it did not reveal how much this particular genome cost, but says by June its materials cost will be down to $1,000 per genome. The company aims to launch commercially that month, sequence 1,000 genomes this year and 20,000 human genomes next year.




ROCHE

A few centres have now signed on for pilot projects in which Complete Genomics will sequence five genomes at $20,000 apiece. Many scientists are reserving judgement until they see the results. "The model of outsourcing for generating complex data sets makes scientists nervous," says Rick Myers, director of the HudsonAlpha Institute for Biotechnology in Huntsville, Alabama. "There's a lot of complexity in everybody's genome, and it will be important to have that right."

The quest for accuracy and speed

An hour before Complete Genomics' presentation, the chief technology officer of its cross-town rival, Pacific Biosciences of Menlo Park, spoke. Stephen Turner unveiled a completed genome of the bacterium *Escherichia coli*, saying the company had covered each base an average of 38 times for an accuracy of greater than 99.9999%.

Pacific Biosciences uses a single-molecule technology with DNA polymerase, the enzyme used by cells to assemble DNA strands, that reads out the product of the sequencing reaction as it progresses. Although its current machines read 3 bases per second, it aims to produce entire human genomes in under three minutes by 2013. It has also promised to deliver longer read lengths; Turner says the average read length of the *E. coli* genome was 586 base pairs, with some as long as 2,805 base pairs — "higher than any other read length in production," he says. Some scientists hope long read lengths will eliminate errors and allow them to see parts of the genome that are difficult to read.

Pacific Biosciences intends to launch commercially late next year. Meanwhile, current sequencing technologies — such as those sold by Illumina, based in San Diego; Applied Biosystems, from Foster City, California; and Roche, based in Basel, Switzerland — are pouring out data at an astonishing rate, delivering


How far and fast can gene-sequencing technologies go?

multiple human genomes' worth of data in a single multi-day run (see chart). That rate continues to increase as prices drop; Illumina, for instance, said at the conference that its technology will be able to sequence human genomes for as little as $10,000 by the end of this year. The current companies, says Nusbaum, are "not going down without a fight".

Not all entrants in the crowded sequencing market are faring well. Third-generation machines made by Helicos Biosciences of Cambridge, Massachusetts, have been dogged by sequencing errors. And just days before the conference began, Helicos revealed that its first customer had returned one of the few machines it has sold. At the meeting, the company said that it had assembled the genome of a *Caenorhabditis elegans* nematode worm. But its troubled history and the high cost of its reagents and machine — just reduced to $999,999, compared with around half a million dollars for other next-generation machines — have eroded the confidence of many scientists.

In a pre-meeting workshop, John McPherson of the Ontario Cancer Institute in Toronto, Canada, summarized the general feeling: "They were a pioneer in single molecule sequencing, but I think they've fallen short of their goals." Yet Helicos' chief technological officer William Efcavitch dismisses the critics. "The rumours of our demise are greatly exaggerated," he said at the conference.

Many scientists hope he is right, and that Helicos and many other companies keep competing to deliver more data at lower prices. "The competition has been really healthy," Myers said. "It seems like a miracle that we can get 80 million sequence reads in a few days now, but no matter how well [the companies] do, we want more." ■

Erika Check Hayden

HAVE YOUR SAY
Comment on any of our News stories, online.
www.nature.com/news

Britain hits a hurdle in replacing key animal-pathogen facility

Plans to build a £121-million (US$180-million) animal-pathogen facility in the United Kingdom are likely to have to be scaled back after the Department for Environment, Food and Rural Affairs (DEFRA) walked away from the project.

The decision is a "big setback for the United Kingdom's ability to respond to animal diseases and to protect the livestock industry", says Keith Gull, a microbiologist at the University of Oxford and chairman of the governing body for the country's Institute for Animal Health (IAH).

In 2005, DEFRA agreed with the Biotechnology and Biological Sciences Research Council (BBSRC) and the Department of Trade and Industry to redevelop ageing facilities at the IAH's Pirbright site — in 2007, a poorly maintained drainpipe there led to an outbreak of foot-and-mouth disease. But the project, which was due to be completed in 2011, has been delayed until 2013. Costs have also been rising; £33 million has already been spent, although it is not clear on what.

The BBSRC is now reassessing the size and scale of the plans. "The vision of the joint venture with DEFRA was to bring together our complementary activities; the BBSRC is now taking this project forward alone," says Matt Goode, a BBSRC spokesman.

DEFRA was to have contributed around £58 million to the new building and to move about 70 staff there from its Veterinary Laboratories Agency in Weybridge. DEFRA would not confirm whether it has abandoned the project, but it says in a statement that it has no plans to stop spending £5.8 million per year on research and surveillance on exotic diseases at the IAH.

Ageing facilities at the Pirbright Laboratory need updating.

G. FULLER/PA

The news comes as the UK government published its response, on 3 February, to a review by microbiologist Iain Anderson of the foot-and-mouth outbreak. The Anderson report recommends redeveloping the IAH into a new National Institute for Infectious Diseases, which would be funded by multiple sources, including government. Such an institute was meant to join up and strengthen government research on diseases affecting animals and humans. At present, animal-health research is overseen by DEFRA and the BBSRC, and human-health research by the Department of Health and the Medical Research Council.

The government's response to the Anderson report notes that the BBSRC will continue to fund the IAH. But it says that the government will shortly be consulting on plans to establish a new body that would operate at arm's length from ministers. This body would "assume all of DEFRA's existing roles and responsibilities in relation to animal health".

In a policy statement on how the United Kingdom should take forward infectious-disease research, the Royal Society says that "a national policy for infectious diseases of both humans and animals is needed". It also notes that "research support and policy for infectious diseases is highly fragmented". By contrast, in the United States, four national centres covering infectious diseases in humans and animals are subsumed under a Coordinating Center for Infectious Diseases, funded by the Department of Health and Human Services.

But now the IAH redevelopment will fall short of such goals, and it is unclear whether the government intends to take forward the idea of a national institute. The BBSRC says it will consider its plans and budget for redeveloping the IAH at an 11 February meeting. It will then submit a revised business plan to the Department for Innovation, Universities and Skills — which has replaced the Department of Trade and Industry — to ask for additional funding.

"We will end up with a nice new facility," says Gull. "But will it be fit for purpose for the nation's needs?" ■

Natasha Gilbert

Obama puts focus on FDA after peanut poisonings

Salmonella outbreak prompts review of US food safety.

When the US president worries aloud about the safety of his 7-year-old daughter's peanut butter sandwich, as Barack Obama did last week, it may be safe to assume that change is at last coming to the government's beleaguered food-safety system.

When he announces, in the same interview, "a complete review" of operations at the Food and Drug Administration (FDA), it may be equally reasonable to expect new things at the $2.3-billion agency responsible for ensuring the safety of most US food and drugs.

Yet no top-to-bottom review will be possible until the agency has a new leader. On 30 January, the White House announced that the appointment of the FDA commissioner was imminent; but that was before the 3 February withdrawal of Tom Daschle, Obama's nominee for secretary of health and human services, to whom the FDA commissioner reports. It is not clear whether Daschle's withdrawal will speed or slow the naming of an FDA chief.

"There has to be institutional change."

Although roughly two-thirds of the FDA's budget goes to regulating drugs and medical devices, the agency is also responsible for ensuring the safety of 80% of US foods — and it is in food safety that Obama has run into the FDA's ongoing problems. Since September, 575 people in 43 states — half of them children — have fallen ill in an outbreak of *Salmonella* poisoning that originated at a peanut plant in Georgia.

The outbreak has highlighted the steady weakening of the FDA's food-safety capabilities. Between 2003 and 2008, for instance, the number of food scientists at the agency's Maryland headquarters shrank by 18% even as the number of food-borne disease outbreaks more than doubled.

Obama himself introduced legislation last year aimed at strengthening and expanding food surveillance technologies. It would also have forced the FDA, the Centers for Disease Control and Prevention in Atlanta, Georgia, and state and local public-health agencies to share data on outbreaks of food poisoning and to coordinate investigations more effectively.

R. FELD/AP

At a 5 February hearing of the Senate agriculture committee, chairman Tom Harkin (Democrat, Iowa) waved a jar of Jif peanut butter, saying: "When Americans can't count on the safety of basic items that go into our children's lunch boxes, then we are in big trouble."

This "is a turning point", says Michael Taylor,

One more step for private Moon mission

A spectrometer meant to fly to Mars on a European mission in 2016 will get to the Moon first. The Dutch team that is building the instrument last week announced it would send a scaled-up version, dubbed MoonShot, to the lunar surface by 2011 with Odyssey Moon, a company headquartered in the Isle of Man, UK.

If it works, the private MoonOne lander and its successors could serve scientists much as a commercial trucking company serves wholesalers, providing a platform to ferry science instruments and other payloads to the lunar surface.

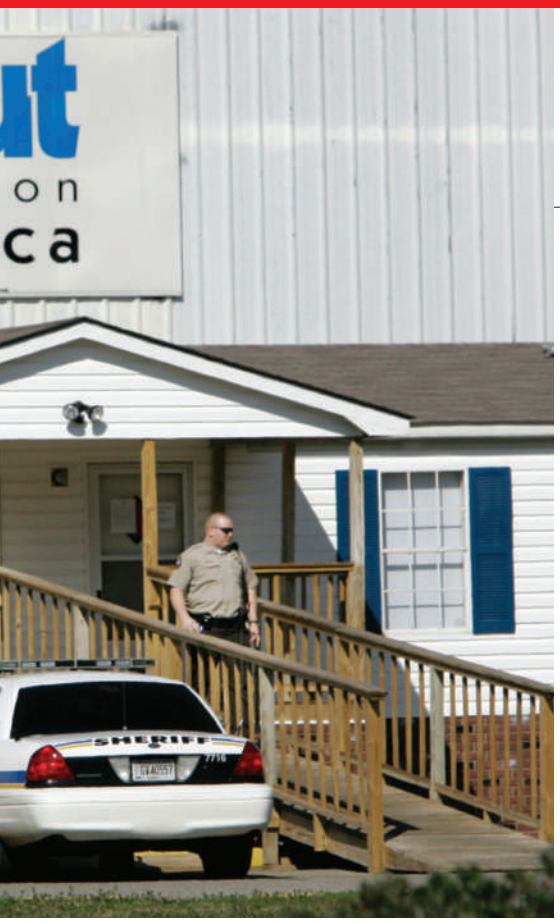
"The intention is to bring on a new age," says Alan Stern, science mission director for Odyssey Moon and former science chief at NASA. That could mean new opportunities for scientists whose instruments were cut or scaled back in government missions, or whose nations do not have their own lunar spacecraft. "It's an ongoing business that will make sense many years in the future," Stern says.

MoonShot was originally designed to look for organic compounds in the Martian soil using two different types of laser-based spectrometry. The version that will go to Mars on the European Space Agency's ExoMars mission has been cut back to include only a Raman spectrometer, which uses a laser beam to detect chemical signatures, including organic compounds, in surface samples. The version flying privately to the Moon, however, will also have a laser-induced breakdown spectrometer, which will scan the lunar soil and can detect heavy metals. "You shine the laser and just find out what's there," says Erik Laan, an engineer who helped build the instrument at TNO Science and Industry in Delft, the Netherlands.

"It's an ongoing business that will make sense many years in the future."

MoonShot would be the first Dutch instrument on the lunar surface. A consortium of European companies, including Philips Applied Technologies, Dutch consultants Space Horizon, the Free University in Amsterdam and the Delft University of Technology, will pitch in to pay Odyssey Moon an unspecified amount — although less than US$10 million — to transport the instrument.

The MoonOne lander is already slated to carry two other commercial payloads. These include a precursor to the International Lunar Observatory, a communications dish



Outbreaks of *Salmonella* have been traced to products from the Peanut Corporation of America.

a former FDA deputy commissioner for policy who is a professor of health policy at George Washington University in Washington DC. "Members on both sides of the aisle are getting that there is a system-wide problem here, that there has to be institutional change."

Importantly, the food industry — which has lost hundreds of millions of dollars in the past two years after recalls of tainted spinach and peppers — is also at the table. Last month, ten industry groups, including the Grocery Manufacturers Association, wrote to House and Senate leaders begging them to "quickly enact food safety reforms" by giving the FDA new powers. A bill introduced on 4 February, by Representative Rosa DeLauro (Democrat, Connecticut) would consolidate the agency's food-safety functions at a new Food Safety Administration, whose chief would report directly to the secretary of health and human services.

The FDA's current lack of food-safety teeth was evident at last week's Senate hearing, at which senators grilled agency officials on the history of inspections at the Peanut Corporation of America in Blakely, Georgia. There, company tests revealed *Salmonella*-tainted products in 12 separate instances beginning in June 2007; the company retested and got a negative result, then shipped the products. At least eight people have died after eating peanut products that originated at the plant.

Stephen Sundlof, director of the FDA's Center for Food Safety and Applied Nutrition, clarified at the hearing that companies are under no obligation to report to the FDA when food tests turn up tainted products; that the agency has no power to require that companies recall products; and that it cannot demand the results of a company test unless it has a reasonable belief that an already-shipped product is tainted.

Although the *Salmonella* outbreak is the first major FDA-related issue for Obama, there will undoubtedly be more. A steady stream of reports in recent years from groups such as the Institute of Medicine, the National Academy of Sciences and the Government Accountability Office has documented the woes of an understaffed, underfunded, demoralized and scientifically deficient agency (see *Nature* **450,** 1143; 2007).

A Government Accountability Office report released last month said that the FDA has been lax in implementing stringent review procedures for some of the riskiest medical devices such as heart valves and pacemakers. And the same month, a group of nine dissident agency scientists wrote to Obama charging that their superiors had targeted them for a criminal investigation, after they complained that agency officials are putting patients at risk by approving high-risk devices that have not been proved to be safe or effective.

Whether it is the politics of devices and drugs or the public outcry over food safety, the new commissioner, once installed, will have plenty on his or her plate. Taylor says that the person will need "the skill to work within the agency well but also the strength to work with the outside world — to manage the slings and arrows that will come when you try to make change". ■

Meredith Wadman

SLOAN DIGITAL SKY SURVEY

ASTRONOMY'S TOP TEN
World's telescopes rated by citation count.
www.nature.com/news

ODYSSEY MOON

The MoonOne lander aims to fly cargo to the Moon for cash.

at the lunar south pole envisioned by a Hawaii-based non-profit organization, and packages from Texas-based company Celestis, which acts as a broker for ferrying cremated remains and other relics to the Moon.

Odyssey Moon has also signed up Paul Spudis, a geologist at the Lunar and Planetary Institute in Houston, Texas, who serves as the company's chief scientist. And it has partnered with NASA to model its lunar lander after the US space agency's common spacecraft bus, which uses a modular design that can be manipulated into either a lander or an orbiter.

Robert Richards, Odyssey Moon's founder, describes it as a logistics company with a business model similar to that of FedEx. "We want to get customer packages to the Moon," he says, "and we don't really care what we use to do it as long as it gets there reliably and cost-effectively." Eventually, such private craft might end up helping to resupply government-initiated efforts, such as the proposed International Lunar Network, a series of stations containing scientific instruments that will be positioned at different places across the lunar surface.

Odyssey Moon and 15 other teams are also vying for the Google Lunar X Prize, a US$20-million race to be the first privately funded spacecraft to relay high-definition video, images and data from a lunar lander back to Earth. The teams vary in their backgrounds, interests and long-term goals. A team named Next Giant Leap, in partnership with the Massachusetts Institute of Technology's Space Systems Laboratory and MicroSat Systems of Littleton, Colorado, among others, is focusing on meeting the prize's minimum criteria and has no plans yet to develop a commercial craft of its own. Astrobiotic Technology, founded by robotics specialist William Whittaker of Carnegie Mellon University in Pittsburgh, Pennsylvania, plans to launch exploratory missions, first to the Apollo 11 site and then to the lunar poles.

Competitors for the Google Lunar X Prize have until 31 December 2014 to stake their claim on the cash awards and the Moon. ■

Ashley Yeager

Neglected disease boost

Fresh funding aims to raise awareness and improve control measures.

The Bill & Melinda Gates Foundation has called for a worldwide push to slash the burden of neglected tropical diseases such as elephantiasis, trachoma and schistosomiasis by 2020.

On 30 January, Bill Gates announced a US$34-million grant to leverage new funding globally to prevent and treat these illnesses, which affect 1.4 billion people. The money will go to an international alliance, the Global Network for Neglected Tropical Diseases, which aims to raise awareness and bring previously fragmented disease-specific efforts under a common umbrella to better coordinate fund-raising and field operations.

The trick, says Gates, is bringing in the substantial new investment needed to drastically scale up control measures using existing tools.

Such scale-up has already proved its mettle: take, for example, lymphatic filariasis, a devastating parasitic disease usually known by its worst manifestation, elephantiasis. Just nine years ago, drugs for lymphatic filariasis reached only 25 million people in 12 countries; more than one billion people in 83 countries are at risk. But a public–private partnership, the Global Programme to Eliminate Lymphatic Filariasis, whose secretariat is at the Liverpool School of Tropical Medicine, UK, is now bringing vital drugs to 570 million people in 48 countries annually. The programme, launched in 2000 and jump-started with a $20-million grant from Gates, estimates that it has prevented the disease in 6.6 million newborns and prevented 2.2 million new serious clinical cases, as well as stopping 9.5 million people already infected from developing severe complications[1].

"The alliance has amply demonstrated proof-of-principle that rapid country scale-up makes a difference," says Alan Fenwick, a parasitologist at Imperial College London and head of the board of the Schistosomiasis Control Initiative.

The Global Network, inaugurated by former President Bill Clinton in 2006, is headquartered at the Sabin Vaccine Institute in Washington DC. With the new Gates money it hopes to drastically step up prevention and treatment of the seven most common neglected diseases that can be prevented and treated with existing drugs for about 50 cents per person annually — trachoma, onchocerciasis, schistosomiasis, lymphatic filariasis, trichuriasis, hookworm and ascariasis.

The key is negotiating donations of billions of tablets from drug companies and purchasing others at cheap rates by negotiating bulk deals at a global scale, at prices far below those that would be attainable by any single country. A 'rapid impact package' — a cocktail of four drugs that targets all seven diseases in a single treatment — is then used for mass preventative chemotherapy across entire populations.

The huge logistics are handled by an innovative community-based approach, made up of far-reaching networks of hundreds of thousands of volunteers in individual communities, and local health services. "You can put thousands of drugs on the back of a bicycle," says David Molyneux, a parasitologist at the Liverpool School of Tropical Medicine and

Dutch university slashes evolution staff

Darwin200

The Year of Darwin has got off to a bad start in the Netherlands, where a national reorganization of university budgets has led Leiden University to sack its classical evolutionary-biology staff.

"There will be no one left who can teach natural selection," says population ecologist Jacques van Alphen, one of half a dozen tenured professors who will lose their jobs on 1 March. Some technicians and postdocs will also be fired.

The scientists are challenging the legality of the dismissals in court, and have launched a petition that has been signed by more than 3,000 researchers worldwide. They claim that their jobs have been eliminated in favour of jobs in molecular biology. Around 30 faculty members will remain on the staff at the university's Institute of Biology.

Like all universities in the Netherlands, Leiden is experiencing the consequences of a decision by science minister Ronald Plasterk, a former molecular geneticist, to introduce greater competition in the scientific community.

"There will be no one left who can teach natural selection."

Last September, the government decided to transfer €100 million (US$130 million) of its budget for universities to the NWO, the Dutch granting agency, for distribution through national competition. The sudden shortfall has meant that universities with little financial buffering, such as Leiden, are having to cut into their own flesh. Leiden found itself with a shortfall of €10 million — €2.5 million of which it passed on to the Institute of Biology.

Mathematician Sjoerd Verduyn Lunel, dean of natural sciences, says the institute considered carefully where it would trim. "We applied the same criteria for identifying where to make the cuts as the government uses to fund the universities," he says. The list of criteria includes factors such as the numbers of students and income from competitive grants.

The molecular biosciences have been


A programme tackling elephantiasis has prevented infection in millions of people.

R. U. CHANDRAN, TDR, WHO/SPL

executive secretary of the lymphatic filariasis effort. He says that the effects of the drugs in the field are "staggering".

Gates hopes that the new grant will help to catalyse substantial additional funds. The alliance will use the grant to launch End the Neglect 2020, a campaign of high-level political advocacy, public fundraising and wooing industry for drug donations. The idea is to push control of some of the most neglected diseases up the international political agenda and raise public awareness, says Fenwick. "The fight against neglected tropical diseases is at a tipping point," adds Peter Hotez, president of the Sabin Vaccine Institute.

The network's past advocacy efforts are already paying off. In 2007, the US Congress gave $100 million through the US Agency for International Development to distribute a rapid-impact drug package in nine countries. Last year, the US government pledged a further $350 million, and the United Kingdom offered £50 million ($75 million), over five years, to control neglected diseases, although these commitments have yet to be disbursed. The pharmaceutical companies Merck, GlaxoSmithKline, Pfizer, Johnson & Johnson and MedPharm have also committed to large donations of drugs up until 2020.

"The advocacy has been a big positive aspect," says Bernard Pécoul, head of the Drugs for Neglected Diseases Initiative based in Geneva. But he criticizes the alliance for focusing on the seven diseases that can be easily treated, largely ignoring others such as sleeping sickness, visceral leishmaniasis and Chagas disease.

Such tensions are fuelled by the sheer shortage of resources. Neglected diseases have long faced utter indifference from funders, with the exception of AIDS, tuberculosis and malaria. These three attract billions; the rest, the 'most neglected diseases', get only a fraction.

The first comprehensive audit of research spending on neglected diseases, published last week in *PLoS Medicine*[2], found a similar picture in research and development to generate new therapeutics, with almost 80% of the $2.5 billion invested in 2007 going to AIDS, tuberculosis and malaria. By contrast, other neglected diseases each picked up from 4.4% to less than 1%.

To ease this, the Geneva-based Global Fund to Fight AIDS, Tuberculosis and Malaria should be balanced to include other neglected diseases, says Molyneux. Not only would this increase resources for the most neglected diseases, but the fund's three current targets could also benefit from a more integrated approach, he argues. Schistosomiasis, he notes, can cause vaginal lesions that treble the risk of HIV infection[3].

"There is little point focusing on stopping people dying from AIDS or malaria if they are going to be disabled by these other diseases."

"There is little bloody point focusing on just stopping people dying from AIDS or malaria," says Molyneux, "if they are then going to be disabled and blinded by these other diseases."

In a paper in the *Lancet* last month[4], Molyneux and his colleagues argue further that the distribution of malaria drugs could be greatly expanded by piggybacking on the alliance's vast and established community-distribution networks. ■

Declan Butler

1. Ottesen, E. A. *et al. PLoS Negl. Trop. Dis.* **2,** e317 (2008).
2. Moran, M. *et al. PLoS Med.* **6,** e1000030 (2009).
3. Kjetland, E. F. *et al. AIDS* **20,** 593–600 (2006).
4. Molyneux, D. H. *et al. Lancet* **373,** 296–297 (2009).

GOT A NEWS TIP?
Send any article ideas for *Nature*'s News section to **newstips@nature.com**

K. CAMPBELL/GETTY

more successful in attracting grants than evolutionary biology at Leiden, Verduyn Lunel says. "It is a sad situation, but if you have a government strategy to increase quality through competition, how else can you implement it? There has been a growing discrepancy in evolutionary biology over the past few years which everyone knew about." Jobs in molecular evolution and ecology have survived the cuts involved in the Leiden decision.

The dismissed evolutionary biologists point out that the University of Groningen is the only other centre in the country that teaches traditional evolutionary biology. "What's happening in Leiden is a real pity," says Serge Daan, dean of science at Groningen, which has been able to absorb most of the cutbacks passed on from the NWO decision. "In biology you need an integrative approach — students need to understand how things operate on the functional, population level as well as the molecular level."


Jacques van Alphen's job is one of those due to go.

J. VAN ALPHEN

Isabelle Olivieri, a population biologist at the University of Montpellier in France and president of the European Society for Evolutionary Biology, sees the fate of the Leiden scientists as a sign of more widely fading political support. She points out that the French National Research Agency has recently substituted a proposed national research programme on general evolutionary biology with one focused on biodiversity and conservation.

"There is little focus on the important middle ground between this and molecular evolution, which by its nature does not bring much money into universities," says Olivieri.

The Royal Netherlands Academy of Arts and Sciences has also voiced its concern about the Leiden decision. ■

Alison Abbott

US nuclear agency to consider defensive move

The powerful budget authority within Barack Obama's administration is considering the consequences of moving the US nuclear-weapons complex from its traditional civilian home within the Department of Energy to the Department of Defense.

In an undated memo, the Office of Management and Budget has asked both departments to submit a report by 30 September 2009 outlining the pros and cons of transferring the $9.1-billion National Nuclear Security Administration (NNSA) — a move that could happen as early as 2011. The quasi-independent agency oversees $6.3 billion in activities at the five facilities and three national laboratories that are responsible for the design, creation and stewardship of nuclear weapons.

Although the agency has been attacked from many corners for its perceived inefficiencies and micromanaging style, congressional supporters of the weapons complex are worried about the potential move's consequences for the non-weapons science that goes on at facilities such as the Los Alamos National Laboratory in New Mexico.

Jude McCartin, a spokeswoman for Senator Jeff Bingaman (Democrat, New Mexico), says, "Senator Bingaman has very serious concerns about narrowing the mission down to just defence work and what that might do to the laboratories."

"People are sick to death of following silly bureaucratic rules."

Other defence-policy experts questioned the defence department's lack of experience with the large scientific facilities managed by some weapons labs, such as the National Ignition Facility, a laser-fusion experiment at the Lawrence Livermore National Laboratory in California. "The labs are 'big science' labs," says Clark Murdock, a senior adviser at the Center for Strategic and International Studies in Washington DC. And the defence department "has never done big science", he says.

By contrast, many in Congress, including Bingaman, are ready to reform the NNSA. It was set up in 1999 after the Wen Ho Lee security scandal at the Los Alamos lab exposed security and management issues at the energy department. A year later, a congressional oversight panel found the agency to be nothing more than a "paper organization, bereft of the leadership, structure, and degree of semi-autonomy intended by the Congress."

J. RAEDLE/NEWSMAKERS/GETTY

Non-weapons work at national labs could suffer.

Paul Robinson, former president of Sandia National Laboratories in New Mexico, says that the agency has become wasteful and excessively risk-averse. "The morale over these [nuclear-weapons] missions is now near zero because people are sick to death [of] following these silly bureaucratic rules," he says. A spokesperson at the NNSA declined to comment.

In December 2006, a task force convened by the Defense Science Board recommended a different approach for managing the nuclear-weapons complex: dissolving the NNSA in favour of an entirely new organization — the National Nuclear Weapons Agency — that would report directly to the president. Murdock says such an agency could pursue its nuclear-weapons mission while the labs receive separate science funding streams from the departments of defence, energy and homeland security.

Any changes are likely to face stiff bureaucratic resistance. After the Department of Energy was created in 1977 from the remains of the Atomic Energy Commission, the administration of Ronald Reagan in the 1980s tried and failed to shift the weapons complex to the defence department. Now, the incoming Energy Secretary, physicist Steven Chu, seems ready for the next showdown. On 22 January, Chu addressed all Department of Energy employees in a video transmission from Washington DC, and was asked what role he saw for the NNSA. "It's going to remain a part of the Department of Energy, how's that?" he replied. ■

Eric Hand

Obama demands rethink of regulatory processes

President Barack Obama is changing the way that the US government sets its regulations. The move is one of a series of actions that will affect science-informed rules from pollution standards to food-safety inspections.

Obama has asked Peter Orszag, the new head of the Office of Management and Budget (OMB), to suggest revisions to the office's fundamental principles for reviewing proposed government regulations. He has also revoked an executive order, signed in January 2007 by notoriously regulation-averse former President Bush, that had allowed a political appointee in each agency to veto proposed regulations before they became public.

Rick Melberth, head of regulatory policy at the Washington DC group OMB Watch, says that analysts there are "very happy" with Obama's early steps to reform. Many of the problems, he says, predate the Bush administration. "For three decades, there have been an increasing number of obstacles placed in the way of agencies being able to regulate in a timely and effective manner."

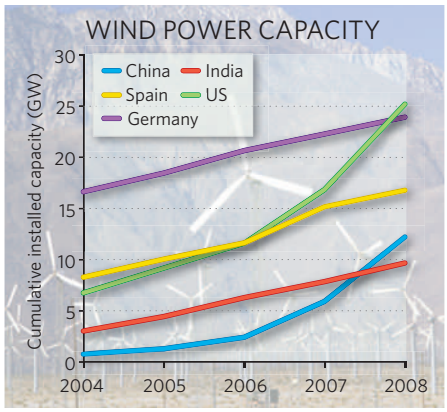
US takes pole position in generation of wind power

The United States has surpassed Germany as the world's largest wind-power producer, according to statistics released by the Global Wind Energy Council earlier this month.

Despite many projects being delayed owing to the economic crunch last autumn, wind producers around the world installed more than 27 gigawatts of capacity in 2008, bringing global capacity to nearly 121 gigawatts (see chart). The United States led the way with 8.4 gigawatts of new capacity, to reach nearly 25.2 gigawatts in total. Wind turbines provided about 42% of the nation's new power-producing capacity in 2008.

Collectively, Europe added 8.9 gigawatts, bringing its total to 66 gigawatts. China more than doubled its capacity, reaching 12.2 gigawatts, and industry officials say that they expect the country's capacity to double again this year.

The council estimates that the industry now employs some 400,000 people, with the 2008 market for new turbines and installations ringing in at $47.5 billion.



SOURCE: GWEC/WWEA; IMAGE: J. SOHM/GETTY IMAGES

Lucy's museum tour threatens to become a spell in storage

The 3.2-million-year-old bones of 'Lucy' (pictured), the skeleton of the *Australopithecus afarensis* hominin discovered in Ethiopia in 1974, have undergone a complete high-resolution computed tomography scan. Researchers at the University of Texas at Austin have thereby created a digitized record that will allow the bones' internal structure to be examined (inset).

But interest from US museums in hosting an exhibition featuring the 40% complete skeleton is lagging, and officials say it will go into indefinite storage next month at the Houston Museum of Natural Science in Texas.

A $2.25-million, five-month Lucy exhibition at the Pacific Science Center in Seattle, Washington, which closes on 8 March, is the second leg of what was originally planned to be a ten-city, six-year tour, although concerns were raised over the effects the tour would have on the delicate fossil (see *Nature* **444,** 8; 2006). But total attendance at the Seattle exhibition is likely to be a third of the 250,000 projected, and no other museum has yet signed up for the exhibition.

F. WHITE/HOUSTON MUS. NATURAL SCI.

J. MAISANO/UNIV. TEXAS, AUSTIN

Pakistan releases nuclear smuggler from house arrest

Abdul Qadeer Khan, the father of Pakistan's nuclear bomb and a central figure in a major nuclear smuggling ring, has been freed.

A brief order issued on 6 February by the Islamabad High Court released the 72-year-old metallurgist from house arrest. Khan had been confined since 2004, after confessing to selling Pakistani nuclear technology to Iran, North Korea and Libya. The government had pardoned him after his admission but kept him under strict guard in his Islamabad home.

Khan's visitors and phone calls will be monitored by the Pakistani Interior Ministry and he will not be allowed to leave the country.

Sweden backs construction of new nuclear plants

The Swedish government has overturned its almost 30-year-old ban on the construction of nuclear power plants.

The moratorium, which dates back to 1980, committed the government to shutting existing plants and banned the building of new ones. Ten reactors still supply the country with nearly half of its electricity.

The new policy, announced by the coalition government on 5 February, annuls the phase-out and allows for the construction of replacements for the ten existing reactors. The government thinks that nuclear power stations will be needed to meet the nation's goal of no net greenhouse-gas emissions by 2050.

Sweden is the second European nation to reverse its nuclear policy in the past year. In May 2008, Italy announced it planned to resume building new nuclear power plants.

Transgenic drug gets green light from the United States

The US Food and Drug Administration (FDA) has issued its first approval of a drug produced by a genetically engineered animal.

The drug, marketed as ATryn, is the human blood protein antithrombin and will be used to treat various blood-clotting disorders. It is produced in the milk of transgenic goats engineered by GTC Biotherapeutics, a company based in Framingham, Massachusetts.

The FDA's decision, released on 6 February, had been delayed while the agency worked out how to regulate products made from genetically engineered animals. It finally issued industry guidelines on 15 January this year (see *Nature* 457, 371; 2009), and it was widely expected that approval of ATryn, which had already been given the green light by a scientific advisory committee, would soon follow. The drug was approved in Europe in June 2006.

THE OTHER STRAND

Geneticists looked to the human genome to understand human evolution. But it's hard to interpret without considering the inheritance of culture, finds **Erika Check Hayden**.

Darwin200

Barely a decade after Charles Darwin published *On the Origin of Species*, he and his long-time correspondent Alfred Russel Wallace were engaged in a fierce debate. Darwin said that natural selection had shaped the human species just like any other. But Wallace disagreed, arguing that selection alone could not account for the exceptional capabilities of the human mind. "How could natural selection, or survival of the fittest in the struggle for existence, at all favour the development of mental powers so entirely removed from the material necessities of savage men?," he wrote[1].

Wallace lost out. By the mid-twentieth century most scientists had agreed that human bodies and minds were the product of genes that had evolved under the pressures of natural selection, just like everything else in the living world. One of the exciting prospects of reading the human genome was that it would reveal the ways in which this had happened — the marks left by evolution as it shaped humans into a species with language, learning and all sorts of other traits peculiarly interesting to it. "The idea was that if we could just identify those few critical genetic differences, we could explain the differences in cognition and language," says Todd Preuss of the Yerkes National Primate Center at Emory University in Atlanta, Georgia.

But today, nearly a decade after the human genome was sequenced, some geneticists are thinking again. Genomics has identified many sequences that are under selection, but it has not provided the simple read-out of human evolutionary history that some had hoped for. Scientists are having to rethink how genomes work, and are now pondering whether genes alone can explain the human animal. They don't think that human biology is incomplete without spirituality, as Wallace did. But they do wonder whether it is incomplete without culture. Because many complex skills and behaviours are being passed on through culture, some researchers are coming around to the view that the species has escaped the need to encode them rigidly in its genome. "Of course the mechanisms of selection are operating," says Ajit Varki, a specialist in human origins at the University of California, San Diego. But perhaps "we don't necessarily fix our behaviours, and we are letting some previously fixed behaviours deteriorate, because we can rely on cultural transmission", he adds.

These ideas are not entirely new. In 1981, geneticists Marcus Feldman and Luca Cavalli-Sforza from Stanford University in California published models to show how human behaviour results from the interaction of biological and cultural evolution. What is new for genome scientists is the realization that they will not be able to interpret the evolutionary marks they have found in the human genome without considering behaviour and environment every step of the way.

Being human

The human species has a unique set of features, including a large brain in proportion to the rest of the body; the ability to communicate complex information through symbolic language; and physiological vulnerabilities to Alzheimer's disease, certain cancers and other conditions. Geneticists hoped to explain the evolution of these human attributes with the tools of comparative genomics — the side-by-side comparison of different species' genomes.

Varki was part of a group that pushed hard to sequence the genome of human's closest relative, the chimpanzee, arguing that any genetic differences between human and chimp sequences would lead straight to the heart of humanness. But the chimp sequence, published more than three years ago[2], hasn't delivered this. One comparison between humans, chimps and mice, for example, showed that the protein-coding sequences of genes expressed

K. LEMON

in the brain have changed very little across species[3].

A few genes are interesting exceptions. Selection seems to have favoured changes in *FOXP2*, a gene involved in human speech, after humans and chimps diverged between 4.6 million and 6.2 million years ago[4]. Geneticist Bruce Lahn from the University of Chicago in Illinois has proposed that *ASPM*[5] and *MCPH1*[6] — both of which are thought to be related to brain size — are under selection in humans. And then there is the gene encoding DUF1220, a "protein domain of unknown function" that seems to be under selection, that is active in the brain and has many more copies in humans than in other species[7]. But none of these genes alone is likely to explain a single human trait. "No real silver bullet has emerged to say, 'This is the human uniqueness gene', and there will never be one such gene," says Evan Eichler, a genome biologist at the University of Washington in Seattle. "It is the collective impact of all these genetic differences that make us human."

Some evolutionary research is leading away from protein-coding genes entirely. In 2006, a team of scientists led by David Haussler at the University of California, Santa Cruz, picked out 49 regions of the human genome[8] that had remained largely untouched throughout the evolution of fish, reptiles, birds and monkeys, and then went into mutational hyperdrive after ancestral humans emerged. The researchers found that the genomic address that has evolved faster than any other codes not for a protein, but for a small piece of RNA — human accelerated region 1 (HAR1) — that is expressed in brain cells during human fetal development. Beyond that, nobody knows what HAR1 does. Its sudden status as a belle at the evolutionary ball underscores the idea that selection could have acted most strongly on sequences outside the bounds of protein-coding genes, leaving Haussler and other researchers to work out what these sequences are doing, and why selection has acted on them. The picture could become even more complicated if, as some recent work has suggested, the statistical tests used to find genes under positive selection are themselves questionable[9]. "There are thousands and thousands of changes to our genomes that have occurred in the past few million years that are still hidden, and the vast majority of those will not be functionally consequential," Haussler says. "So it is an amazingly difficult, needle-in-a-haystack type of search."

"The idea was that if we could just identify those few critical genetic differences, we could explain the differences in cognition and language."
— Todd Preuss

Genome researchers once expected that most of the genetic differences between humans would be in single letters of DNA, a type of variation called a single nucleotide polymorphism, or SNP. But in the past few years, they have discovered that large chunks of the genome can be duplicated, deleted and otherwise rearranged differently between individuals. In a paper published this week[10], Eichler's team analysed the genomes of humans, chimpanzees, orangutans and macaques and found that a burst of duplications appears to have occurred in the last common ancestor of humans and chimpanzees. The question now is why? Such rearrangements risk disrupting essential genes and have been linked to human diseases such as autism and schizophrenia.

Varki and Eichler suggest that structural variations may also confer benefits by expanding the range of genetic diversity. The negative side effects might be outweighed by the advantages conferred by new genes or other beneficial arrangements. And the human genome might have been able to tolerate some of the potentially toxic variants thanks to clothing, tools, agriculture and other cultural innovations that allow individuals with these variants to survive. "By allowing individuals to be buffered against natural selection, perhaps culture allows a wider spectrum of genetic diversity to creep in," Eichler says. "Maybe the wider spectrum of diversity allows for more savants and autistic people in the same population."

Researchers have found that the human genome has accumulated more than its fair share of other potentially harmful genetic changes too — in protein coding regions, promoters and even the loss of entire genes. One explanation is that it is a remnant of the frequent population 'bottlenecks' in human history, in which small groups that migrated to new areas established new populations that all carried the founders' mutations. But another possibility, says statistical geneticist Gilean McVean at the University of Oxford, UK, is that the human ability to learn and adapt has eased the selection pressure that would weed

out some of these changes. "When you look in the human genome, one of the things you see is that it has accumulated a lot of apparently bad mutations, and to some extent humans' inventive skills might have allowed that," he says. A modern example, he points out, might be the ability of humans to make spectacles to counteract poor vision.

Anthropologist and neuroscientist Terrence Deacon of the University of California, Berkeley, has long argued that culture could have "relaxed" human selection. Apes, for instance, have lost the ability to make vitamin C because the gene that facilitates this process has broken down. Deacon suggests that the availability of fruit eased the selection pressures working against individuals who couldn't make their own vitamin C, allowing the relevant gene to accumulate mutations but also making apes dependent on external sources for the nutrient. The relaxed selection created by human culture similarly could have allowed the evolution of more diversity and complexity, Deacon says, but it has also made humans more reliant on the innovations that freed them from selection in the first place. "We have produced symbolic communication and culture and technology, all of which play a part in shielding us from certain kinds of selective forces," he says.

DLILLC/CORBIS

M. SHAH / GETTY IMAGES

Comparisons between the genomes of chimps and humans have revealed some sequences that are under selection.

Genetic burst

When scrutinizing the genome, some researchers see more dramatic evidence of culture's influence. In 2007, a team led by anthropologist John Hawks of the University of Wisconsin, Madison, and genome scientist Robert Moyzis of the University of California, Irvine, proposed that culture is hastening human evolution[11]. The team combed through a set of 3.9 million SNPs from European, African and other ancestral human populations, looking for those that bore a signature of positive selection: they were relatively young, common in individuals from the same population and different from those in the other populations. Such mutations are probably being 'swept' to abundance shortly after appearing because they lie in or near pieces of DNA that are beneficial in some way. The team found that a burst of these SNPs had appeared about 10,000 years ago, suggesting that selection has been causing very rapid genetic change since that time.

> **"By allowing individuals to be buffered against natural selection, perhaps culture allows a wider spectrum of genetic diversity to creep in."**
> — Evan Eichler

The researchers then tested whether evolution had sped along like this throughout human history. If so, then the different ancestral populations should have fairly similar genomes because many mutations and their surrounding regions would have swept throughout the human genome before these populations diverged. But current human populations are much more genetically diverse than this hypothesis predicts, so Moyzis and Hawks have concluded that evolution must have ramped up over the past 40,000 years. They chalk some of this acceleration up to human population growth, which exposed the species to more new mutations and created more raw material for selection. But the other reason, Hawks thinks, is culture — because although the physiology of humans has not changed much in the past 40,000 years, their expansion and migration means that lifestyles, languages and technologies certainly have.

Although not everyone agrees with Hawks's claims, the best understood example of recent human evolution does seem to fit. Genetic mutations that allow adults to digest lactose, a sugar found in milk, have emerged independently in different populations in response to the same cultural innovation — cattle domestication[12]. "I don't see culture as an alternative to genetics, I see culture as being the explanatory factor for these genetic changes," says Hawks. "There is no explanation for change without the gene–environment interaction."

Hawks and others are now looking for the beneficial sequences connected with the selected SNPs. But these sequences are hard to pinpoint, and even harder to connect up with a specific human trait. "We know that genes involved in immunity, genes involved in brains, these are showing signals of adaptation in a way that suggests [these traits] have been important," says McVean. "But that's kind of obvious — you only have to look at our biology or our physiology or behaviour to see that."

To Varki and others, the absence of sequences that underlie specific human behaviours could itself testify to the importance of human culture. According to the 'Baldwin effect', named after the American psychologist who proposed it in the late 1800s, behaviours crucial to survival will often become 'hard-wired' into the genome to ensure that they are not lost. But most human skills are not hard-wired: people who have never lived in the Arctic would have a difficult time figuring out how to hunt a seal,

P.-A. PETTERSSON/GETTY IMAGES ; ATLANTIDE PHOTOTRAVEL/CORBIS

Agriculture, tools and other cultural innovations may have allowed mutations to accumulate in the human genome.

skin it and stretch its hide to build a kayak. But if they moved there, they could quickly learn.

The human species has spent most of its history wandering through and creating new environments, and specific skills or resources can quickly become obsolete. So it might have been helpful to cement into DNA the social and intellectual capacities to learn, but leave out the specific instructions for building a kayak. "It is good to make learning more rapid and to improve memory," says geneticist Eva Jablonka of Tel Aviv University in Israel, "but it is not good to specify it too precisely because this memory and learning allows you to cope with a wider range of environments for which you cannot be prepared genetically."

If genes and culture are evolving together, then some difficult questions start to arise — in particular, about whether human evolution is dividing the species. Surveys of human diversity have shown that people can be classified into genetically similar groups that correlate with their geographic ancestry. And Jonathan Haidt of the University of Virginia in Charlottesville has proposed that genetic differences between these groups could underlie variations in traits such as aggression, thriftiness and spontaneity. "Each [human] niche has its own microclimate with its own adaptive pressures," says Haidt, who was impressed by Hawks' work. "The discovery of much more rapid genetic evolution leads directly to the prediction that we will find dozens or hundreds of genetically based ethnic divergences in traits, many of which will have some moral significance." No such ethnically linked divergence has been found to exist, but Haidt points to a hypothesis that selection has boosted the average IQ of the Ashkenazi Jewish population because of its historical reproductive isolation and history of working in the financial trades[13].

Altered state

Most geneticists disagree with Haidt. They point out that occupations, conflicts and other aspects of culture have changed so rapidly during human history that the selective pressures associated with them would have not had time to fix changes in the genome. And Jablonka returns to her point that it is more evolutionarily advantageous for the human species to stay adaptable than to cement personality traits in the genome of different groups. "It would not only be detrimental," she says, "it cannot happen in a species that is all the time having to adapt to the changes it is creating in the world."

Until researchers understand the evolutionary basis for human behaviour it will be easy to argue over such ideas but hard to rule them out. So geneticists are moving away from evolutionary just-so stories that are based on single genes and embracing the complexity of the cultural and biological contexts in which humans and their genes operate. Daniel Geschwind of the University of California, Los Angeles, for instance, is analysing how genes interact with each other in complex networks. If one gene in the brain occupies a central node in the human network but a peripheral one in chimps, then it is a strong hint that the gene's function has changed in some meaningful way too. This type of approach should help solve the problem that researchers have run up against repeatedly in genomic analyses: how to determine which human characteristic, or phenotype, is affected by a genetic sequence under selection.

Cognitive psychologists have found in recent decades that some phenotypes once thought to be uniquely human — such as the development of moral codes, or the ability to recognize that other individuals have minds of their own — are seen in other animals. And work by Michael Tomasello of the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, has suggested that human-specific phenotypes may be very subtle. His group found that young chimps, orangutans and toddlers performed about the same on tests designed to assess 'physical' intelligence, such as retrieving a piece of food that had been hidden under a cup and then moved. But the children were better at 'social' intelligence tasks, such as copying an experimenter's actions to retrieve a treat stuck in a tube[14].

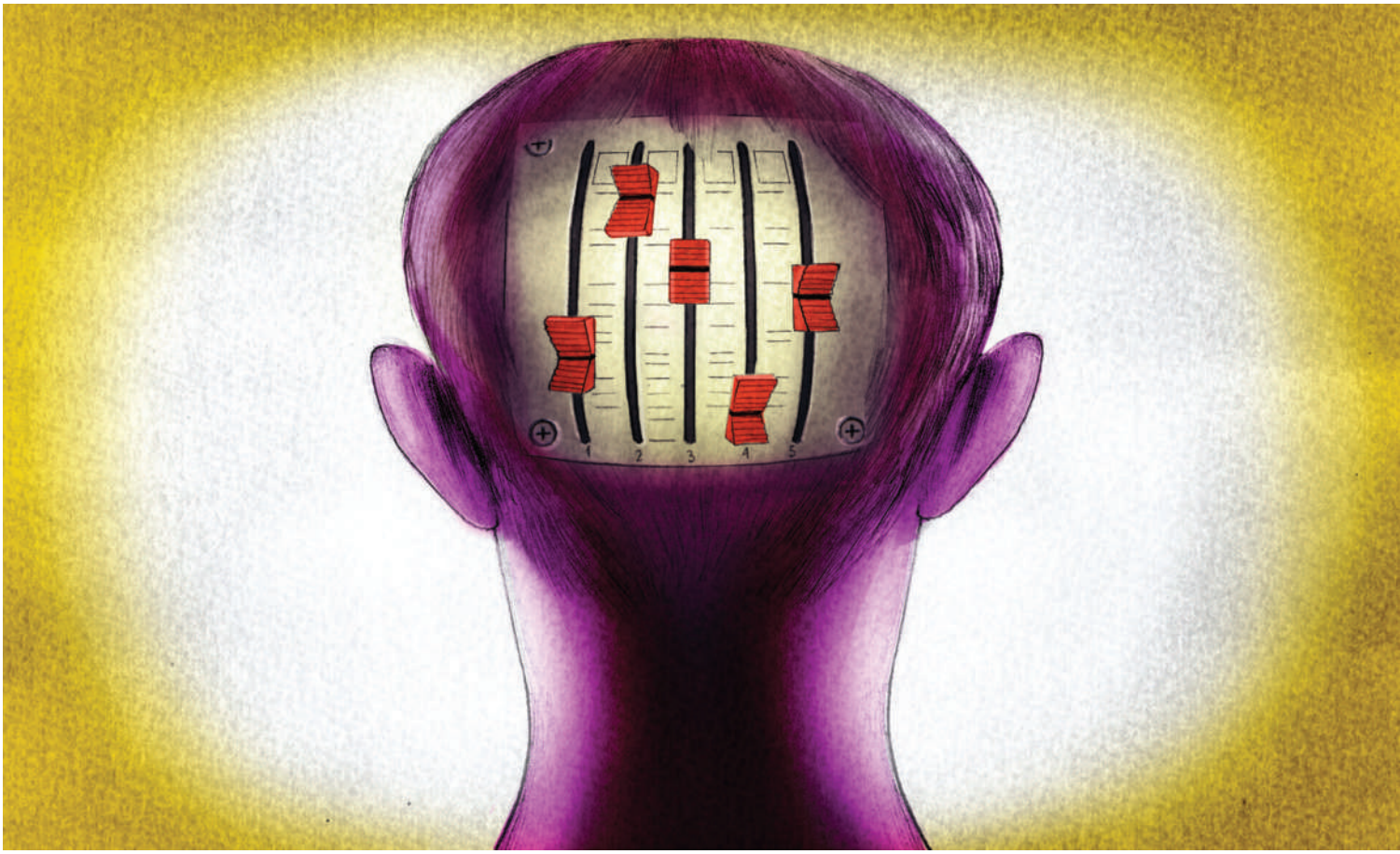
Preuss says that such precise dissections of human-specific traits are still quite rare. "If you go beyond the bland expression of 'advanced cognition' and try to talk about cognitive mechanisms and abilities, we don't really know that much," he says. This means that there is a glut of genomic data but a paucity of crucial information from other fields that would help to make sense of it. "We need to start connecting this genetic world to the traditional anthropological approaches," agrees Hawks, who sees genomics as an inspiration to start collecting and sharing data on an equivalent scale in his own discipline.

Long before the next centennial of Darwin's birth, these data might have closed the book on human evolution. If they do show that culture has shaped the evolution of humans in a way that has no counterpart elsewhere in the animal kingdom, then perhaps Wallace will earn some posthumous credit: this was more than natural-selection-as-usual. But culture cannot have had so strong a role in human evolution without itself being influenced by the results. In the words of Wallace, understanding this interplay will require all of the "mental powers so entirely removed from the material necessities of savage men". ■

Erika Check Hayden is a senior reporter for *Nature* based in San Francisco.

1. Wallace, A. R. in *Contributions to the Theory of Natural Selection. A Series of Essays* (Macmillan, 1870).
2. The Chimpanzee Sequencing and Analysis Consortium *Nature* **437,** 69–87 (2005).
3. Clark, A. G. *et al. Science* **302,** 1960–1963 (2003).
4. Enard, W. *et al. Nature* **418,** 869–872 (2002).
5. Mekel-Bobrov, N. *et al. Science* **309,** 1720–1722 (2005).
6. Evans, P. D. *et al. Science* **309,** 1717–1720 (2005).
7. Popesco, M. C. *et al. Science* **313,** 1304–1307 (2006).
8. Pollard, K. S. *et al. Nature* **443,** 167–172 (2006).
9. *Nature* **457,** 543 (2009).
10. Marques-Bonet, T. *et al. Nature* **457,** 877–881 (2009).
11. Hawks, J. *et al. Proc. Natl Acad. Sci. USA* **104,** 20753–20758 (2007).
12. *Nature* **444,** 994–996 (2006).
13. Cochran, G., Hardy, J. & Harpending, H. *J. Biosoc. Sci.* **38,** 659–693 (2006).
14. Herrmann, E., Call, J., Hernàndez-Lloreda, M. V., Hare, B. & Tomasello, M. *Science* **317,** 1360–1366 (2007).

See Editorial, page 763, and online at www.nature.com/darwin.



Human nature: the remix

People's mindsets are neither fixed by evolution nor infinitely malleable by culture. **Dan Jones** looks for the similarities that underlie the diversity of human nature.

Darwin**200**

Darwin famously gave scant attention to humans in *On the Origin of Species*, contenting himself with the teasing pledge that "Light will be thrown on the origin of man and his history". The promised light came in the *The Descent of Man, and Selection in Relation to Sex* (1871) and *The Expression of the Emotions in Man and Animals* (1872), in which the notion of a common human origin was crucial. In *The Expression*, to further his case that humans shared a great deal of their nature, Darwin "endeavoured to show in considerable detail that all the chief [emotional] expressions exhibited by man are the same throughout the world".

Darwin often relied on anecdotal accounts of travellers and his own casual observations in drawing these conclusions. His poorly sourced claims came, in time, to be challenged by anthropologists focused on the particularities of, and differences between, various cultures. In the 1960s, guided by the prevailing anthropological orthodoxy, Paul Ekman, now retired, set out to prove Darwin wrong by asking for interpretations of facial expressions from the farthest flung people he could get to. He ended up confirming that Darwin had a point. "The evidence is very strong, from studies of both recognition and expression in Western and Eastern, literate and preliterate, cultures that Darwin was indeed prescient," says Ekman. "At least six, perhaps seven, emotions have a pancultural facial expression."

The expression of these emotions is not the only human commonality revealed by cross-cultural studies. People everywhere form communities and pay attention to kinship systems, use complex languages to communicate, socialize and adorn their bodies. Some of these common motifs simply reflect givens of human existence. Many languages use a word related in meaning to 'small person' to describe the opening in the eye's iris, as English does, but this needs no explanation beyond the fact that pupils show the viewer a small reflection of his or herself. Others have deeper significance.

In 1991, anthropologist Donald Brown, then at the University of California, Santa Barbara, published *Human Universals*[1], a survey of hundreds of candidate universal similarities from domains as diverse as language and status

ILLUSTRATIONS BY JONATHAN BURTON

systems to concepts of time and incest taboos. "It was primarily designed to disabuse anthropologists and others of the notion that universals are few and trivial," says Brown, and it immediately found a warm reception among the then-emerging field of evolutionary psychology, which championed the idea of a universal human nature explained by evolution.

But although some anthropologists had simply denied the existence of the human universals that evolutionary psychology craved, others had put forth a subtler view. As the late Clifford Geertz wrote:

> The notion that the essence of what it means to be human is most clearly revealed in those features of human culture that are universal rather than in those that are distinctive to this people or that is a prejudice that we are not obliged to share ... It may be in the cultural particulars of people — in their oddities — that some of the most instructive revelations of what it is to be generically human are to be found[2].

"At least six emotions have pan-cultural facial expression."
— Paul Ekman

Today's research is taking that message, as well as Brown's, to heart. An emerging cadre of anthropologists and evolutionary theorists is taking a new look at the way oddities and difference shed light on the universal. Human nature, this line of thinking holds, is like the control panel on a mixing desk. It is not infinitely malleable: it has its fixed channels and its presets. But it has lots of faders and switches to play around with, and their cultural twiddling can produce a surprisingly wide range of effects.

Is fair fair?

Trying to understand human nature this way has required some academic reconciliation. "We've had to demonstrate to those who emphasize cultural diversity that there is still a useful and interesting way to talk about human nature underpinning the diversity," says Justin Barrett, a cognitive psychologist at the University of Oxford, UK. "At the same time, we've also had to convince those on the biological side that explaining the complexity of human cultural behaviour requires more than a simple application of the rules of animal behaviour." And the reconciliation is still a work in progress. As anthropologist Joe Henrich of the University of British Columbia in Vancouver, Canada, puts it, "We're all comfortable with evolved aspects of the human mind as well as the importance and power of cultural transmission — but there are large swathes of academia that haven't yet reached this constructive consensus."

Henrich's commitment to the cultural plays out in the way he does research. He and his team have gone out and sampled humanity's cultural diversity in various far-flung reaches of the world, combining in-depth ethnographic work with experimental tools from the psychology lab. "A lot of us are trying to be seriously quantitative about these things," says Henrich.

A notable example of the approach comes from a staple of behavioural economics, the 'ultimatum game': one player of a pair is given a sum of money, say ¥100, and has to offer the other some proportion of it. If the offer is accepted, the money is split as proposed; if it is rejected neither player gets a thing. The proposer thus has to gauge what his or her partner will think fair, or at least accept. Western subjects often reject offers of less than 30% of the money. In a landmark study, Henrich's team took the game on a world tour to 15 diverse small-scale societies. In some places people rarely or never rejected very low offers; in others, surprisingly, they rejected more-than-fair offers of greater than 50%[3]. This fits with the finding that 'altruistic punishment' — people's willingness to punish free-riders who parasitize the efforts of others — varies dramatically across societies[4], even while being present in some degree in most. The fairness fader, if there is one, can clearly be set at a wide range of positions.

Enter the WEIRDos

These studies call into question conclusions about universals derived from looking at limited, and possibly unrepresentative, groups of subjects — which is a problem with most psychology studies. Henrich classifies them as WEIRD — Western, educated, industrialized, rich and democratic. What's worse, undergraduates, the lab rats of psychological studies, are weird even by the standards of WEIRDos.

Why should this make a difference? "Our kids grow up with a lot of active teaching compared with small-scale societies," says Henrich. "Our brains are trained for the particular and strange world we inhabit, one where we're not foraging for food, hunting game or constantly under the threat of disease." Although WEIRDos may use universal foundations to build up their strange ideas (see 'Universal maths'), not everyone has a culture that encourages such things.

The argument is borne out by research: an as-yet unpublished review by Henrich and his colleagues Steve Heine and Ara Norenzayan that looks at available cross-cultural studies confirms that WEIRDos are outliers in many ways. Humans, for instance, have a tendency to prefer a smaller immediate gain to a larger but delayed reward. Yet to say that discounting

Universal maths

There is evidence that humans have not just one innate number system, but two. The 'analogue estimation' system evaluates quantities in an approximate manner by relating them to imprecise notions such as 'amount of stuff' or 'extent of imaginary line'. The second system is more exact, but can initially, and innately, keep track of only three or four items[11].

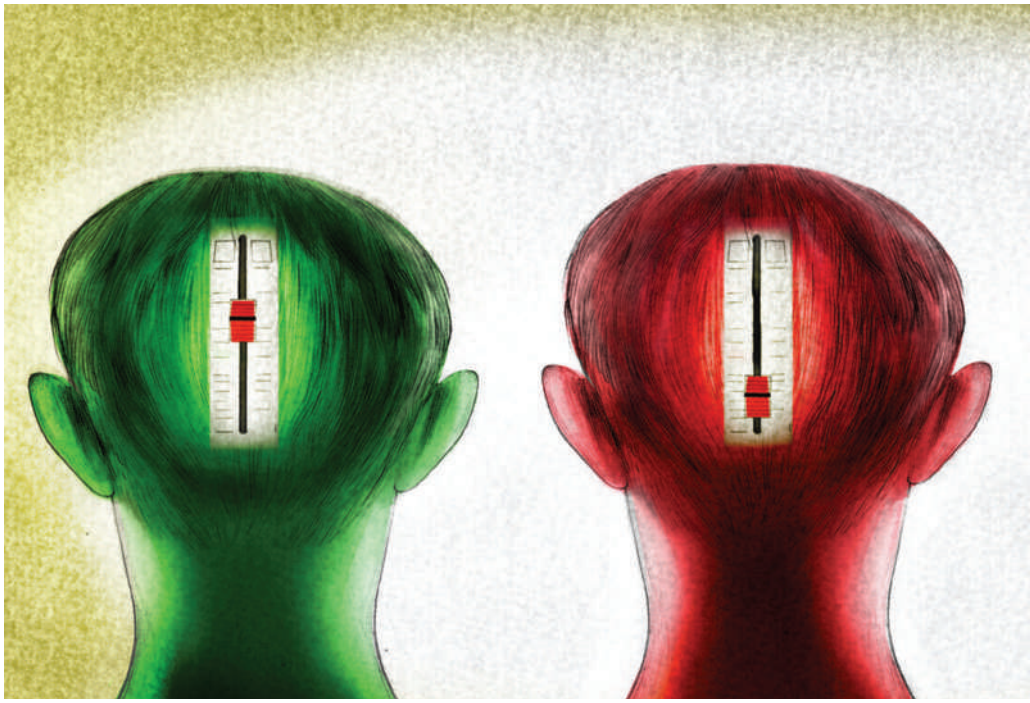
The exact number system can handle larger numbers, but this requires cultural learning and elaboration (the analogue system seems to function in much the same way across cultures). Some societies, such as the Pirahã of the Amazon, have little need for large exact numbers (they do not trade or keep accounts), and use only 'one', 'two' and 'many'. The Pirahã also perform poorly at lining up groups of three or more objects (nine nuts, say, with nine batteries), or estimating when the last of 15 beans has been removed from a tin. Without the cultural demand for trading in large numbers, they have neither developed the concept nor invented the words[12].

Cross-cultural studies also suggest that a basic grasp of geometric concepts, such as points, lines and parallelism, is a universal component of human cognition, or at least found in disparate cultures[13]. A similar story holds for how people relate numbers to space, but with an interesting twist. Western adults tend to order clusters of dots (grouped in 10, 100 and 1,000), along a linear scale, so that the distance between the group of 10 and 100 is much smaller than the gap between 100 and 1,000. Western infants, by contrast, group numbers on a logarithmic scale, with the gap between 100 and 1000 the same as that between 10 and 100. Cross-cultural studies among the indigenous Amazonian Mundurucú revealed that both children and adults order numbers logarithmically. Together with the results from Western infants, this suggests that the logarithmic system could be an innate and universal aspect of mathematical thinking, one that gets tweaked in Western populations through formal education in certain mathematical tools and techniques to produce a switch to the linear system[14]. **D.J.**

in this way is a human universal glosses over the dramatic differences in discounting seen in different cultures. Among the Tsimane of the Bolivian Amazon, future discounting is ten times what you find in the United States; by this measure the Tsimane care even less about the future than do American drug addicts, typically regarded as pathological future discounters.

To try to get deeper into these issues, researchers are looking for ways to gauge not what people say, but how they actually think. "Cognition bridges the biological and the cultural," says Barrett. A cognitive approach focuses attention on what human minds come equipped with, and how they develop — a process that involves much more than just unfolding according to some pre-specified blueprint. "What you really want to study are developmental processes, and how they interact with culture," says Henrich. A model here is Noam Chomsky's pioneering work in linguistics (see 'Universal language'). Chomsky, at the Massachusetts Institute of Technology in Cambridge, united the study of universals and diversity across the world's languages by proposing an innate mental system to acquire, use and comprehend language, which operates on the basis of a great deal of culturally specific input. The constraints placed on humanity's vast linguistic diversity by a shared 'universal grammar' does not make the study of languages less interesting — it just makes it more tractable. The same might apply to phenomena as seemingly complex and socially mediated as religion (see 'Universal religion').

Ideas with a distinct Chomskyan flavour have been a stimulus to recent thinking about morality. What counts as a moral transgression, and how one should react to the transgressor, vary from culture to culture. But deeper patterns seem to lurk beneath this surface diversity.



Following Chomsky's lead, a number of researchers are working on the idea that an innate and universal moral grammar might underlie human ethical judgements. A series of web-based studies led by Marc Hauser of Harvard University have suggested that moral judgements can be explained in terms of such universal and fundamental moral principles[5,6]. Harm caused by direct physical contact, for instance, is generally deemed to be morally worse than harm arising as a side effect, as are harms caused by specific actions rather than omissions.

But these are early days in fleshing out the tool kit of putative moral principles and parameters. "By the time Chomsky started his work in the 1950s he already had a massive amount of descriptive linguistics from all over the world to play with," says Hauser. "In the case of morality, we don't have anything like what the linguists had 50 years ago. We don't know whether the distinctions we're making are at the right level of abstraction, or whether they are principles or parameters."

Web-based surveys are necessarily biased

Universal religion

Whereas human individuals can be atheists, and an increasing number are, human societies are historically universally religious. The detailed beliefs that people have about the powers and roles of deities are not universal or innate. But that doesn't mean that they don't share a common cognitive basis. "There is no God part of the brain," says anthropologist Scott Atran of the University of Michigan in Ann Arbor. "But there are many different kinds of universal process involved in religious thinking that converge to generate a set of family resemblances between religions across cultures."

Much of the evidence for this claim comes from developmental studies with children. "Supernatural thought is so ubiquitous in part because it is so readily accommodated by human cognitive systems from early childhood," says Justin Barrett, a cognitive psychologist at the University of Oxford, UK. "Children seem naturally predisposed to accept the idea that there exist supernatural intentional agents, that the natural world is intentionally designed and purposeful, and that something of human identity is separable from human bodies, allowing for ready belief in some kind of afterlife." But although supernatural beliefs may be in this sense 'natural', they also typically violate assumptions about how the natural world works that seem also to be hard-wired. A god, for example, has most of the properties of a human, with key exceptions; it may not, for example, have a body.

Atran suggests that such violations can make certain supernatural concepts more salient and memorable. At the same time, too many violations can make them confusing. So the deities and supernatural agents that populate the world's belief systems fall into a family, with each member being like a natural object in many respects, and decidedly unlike anything natural in a few others.

Atran suggests that religious belief may have been particularly important in the cultural evolution of large-scale societies. As societies grow, it can be harder to enforce moral and altruistic norms, and punish free-riders on the public good. This in turn can make such societies less cohesive and less able to compete with other expanding societies. Moral deities do lots of things," says Atran. "Crucially, they define the sacred boundaries of societies and the taboo things you can't do. If you really believe in these moral gods then the problem of punishment becomes easier, as you punish yourself." **D.J.**

towards the more educated and affluent, so the answer is to go and look at less WEIRD data — which makes the picture more complex. In a recent study[7], Hauser and his Harvard colleague Linda Abarbanell ran a variety of moral scenarios past a largely uneducated rural Mayan population. Compared with web-based subjects, this decidedly non-WEIRD group gave much less moral weight to distinction between acts and omissions. Hauser and Abarbanell think that the weight put on that dichotomy is a parameter that can be set by local culture, just as some grammatical practices are.

A bird's-eye perspective on moral diversity and uniformity comes from psychologists Jonathan Haidt, of the University of Virginia in Charlottesville and Craig Joseph, of Northwestern University in Evanston, Illinois. Surveying anthropology and evolutionary psychology, they argue that evolution has built into the human mind a preparedness to care about five sets of social issues: fairness and justice; avoiding harm to and caring for others; in-group loyalty; social hierarchy and respect for authority; and the domain of divinity and purity, both bodily and spiritual[8].

"Morality is a social construction, but each society constructs it on top of these five innate moral foundations, relying on them to varying degrees," says Haidt. "Some moralities, such as those of secular Europe, rest primarily on the first two, prizing concerns about harm and fairness above all else; other cultures, such as those of traditional India, emphasize fairness less, and the virtues of respect and spiritual purity more."

"Supernatural thought is so ubiquitous because it is readily accommodated by human cognitive systems."
— Justin Barrett

Liberal differences

Haidt and his graduate student Jesse Graham also argue that differential use of the five foundations illuminates the difference between American liberals and conservatives: whereas the former have hypertrophied sensitivity to fairness/justice and harm/care, and atrophied interest in the rest, conservatives are sensitive to all five channels[9]. This inevitably leads to disagreements about what counts as a moral issue, and differences in how moral debates are approached.

Moral norms are a prominent, but by no means exclusive, subset of the broad class of cultural norms. These are the social rules that govern a great deal of people's behaviour in everyday life — from what to wear and how to meet and greet people, to tipping and queuing. Although invoked remarkably often, the psychology of normative behaviour still remains poorly understood.

Psychiatrist Chandra Sekhar Sripada, at the University of Michigan in Ann Arbor, and philosopher Stephen Stich, of Rutgers University in New Brunswick, New Jersey, have surveyed the widely spread literature on normative psychology, and argue that humans are innately and universally equipped with two key pieces of psychological kit that underlie the phenomenon[10]. The first is a norm-acquisition device, in some ways like Chomsky's language acquisition device, which is responsible for picking out and absorbing local norms. The second generates an intrinsic motivation to comply with these acquired cultural norms, which change over time and diverge between groups. If Sripada and Stich are on the right track, then humans are biologically built to be cultural — and different. The universal design specification is not merely compatible with cultural diversity; it is one of the engines that drives it. ■

Dan Jones is a freelance writer in Brighton, UK.

Universal language

Children are prodigious language learners, and natural-born chatterboxes. In the 1950s and 1960s, though, it was widely held that language learning was simply a learning of associations between words and things. Noam Chomsky, a linguistics expert at the Massachusetts Institute of Technology in Cambridge, argued that something more specific was going on. There was, he said, simply not enough data in the child's linguistic environment for an associationist account to be plausible. And the way children speak reflects the use of rules or principles that they could not have inferred from what they had heard. Something must guide the way language develops.

The solution Chomsky provided — an innate universal grammar — remains controversial. The basic idea is that humans are born with a set of mental principles making up a grammar that constrains how any human language can be built. The principles of this universal grammar shape language acquisition and underscore the commonalities seen among the world's languages. But Chomsky's approach also accounts for diversity with the idea that the universal grammar came with 'parameter switches' that are set differently in different languages. Creating an inventory of these principles and parameters — what is fixed and what is flexible — has been a major preoccupation of numerous linguists over the past 50 years.

To take a common example, some languages, such as English, are said to be 'head first' and verbs precede direct objects, which means I would write John ate toast. Japanese, by contrast, is head-last — so Kenji sushi ate. In both languages, the same principle constrains how certain kinds of word or phrases can be put together — the phrase has to have a head and some other words in any order — but the head-position parameter (first or last) is set by cultural learning. Universality and diversity blend together seamlessly. **D.J.**

1. Brown, D. E. *Human Universals* (McGraw-Hill, 1991).
2. Geertz, C. The Impact of the Concept of Culture on the Concept of Man in *Man in Adaptation* (ed. Cohen, Y. A.) (Aldine Publishing, 1968).
3. Henrich, J. *et al. Am. Econ. Rev.* **91,** 73–79 (2001).
4. Henrich, J. *et al. Science* **312,** 1767–1770 (2006).
5. Cushman, F., Young, L. & Hauser, M. *Psychol. Sci.* **17,** 1082–1089 (2006).
6. Hauser, M. *et al. Mind Lang.* **22,** 1–21 (2007).
7. Hauser, M. & Abarbanell, L. *Cognition* (in the press).
8. Haidt, J. & Joseph, C. in *The Innate Mind* Vol. 3 *Foundations and the Future* (eds Carruthers, P., Laurence, S. & Stich, S.) 367–391 (Oxford Univ. Press, 2007).
9. Haidt, J. & Graham, J. *Soc. Justice Res.* **20,** 98–116 (2007).
10. Sripada, C. S. & Stich, S. in *The Innate Mind* Vol. 2 *Culture and Cognition* (eds Carruthers, P., Laurence, S. & Stich, S.) 280–301 (Oxford Univ. Press, 2006).
11. Pinker, S. *The Stuff of Thought* (Penguin, 2007).
12. Gordon, P. *Science* **306,** 469–499 (2004).
13. Dehaene, S., Izard, V., Pica, P. & Spelke, E. *Science* **311,** 381–384 (2006).
14. Dehaene, S., Izard, V., Spelke, E. & Pica, P. *Science* **320,** 1217–1220 (2008).

See Editorial, page 763, and online at www.nature.com/darwin.

CORRESPONDENCE

Multiple metrics required to measure research performance

SIR — Your Editorial 'Experts still needed' (*Nature* **457,** 7–8; 2009) is correct in that no metric alone can substitute for expert evaluation, because no single metric (including citation counts) is correlated strongly enough with expert judgements for it to take their place. But some individual metrics, such as citation counts, are nevertheless significantly correlated with expert judgements. It is likely that a battery of multiple metrics, when considered jointly, will be even more strongly correlated.

The UK Research Assessment Exercise (RAE) provides such an opportunity, alongside the wealth of potential performance indicators that are increasingly available online. Both enable a candidate battery of metrics — such as citations, co-citations, downloads, tags and growth/decay metrics — to be systematically validated against expert judgements, field by field. The 2008 RAE has also provided data that make it possible to do this validation exercise now, across all disciplines, on an important nationwide scale.

Stevan Harnad Institut des Sciences Cognitives, Université du Québec à Montréal, Montréal, Québec H3C 3P8, Canada; and School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK
e-mail: harnad@ecs.soton.ac.uk

Evolution shapes systems, not just genes

SIR — In my book, *Freaks of Nature: What Anomalies Tell Us About Development and Evolution*, I argue that our understanding of biodiversity and evolutionary change is enhanced by looking beyond mutations and population genetics to consider the mechanisms, constraints and biases of development. Jerry Coyne's review of my book (*Nature* **457,** 382–383; 2009) makes it clear that he disagrees.

Coyne argues that his perspective has been fully tested and accepted as the orthodox working model of evolution. Accordingly, he thinks that development may have little to teach us about evolution.

His sharpest criticism is that I have "short-changed orthodoxy". However, over the years many committed evolutionists — including William Bateson, Gavin de Beer, Patrick Bateson, Gilbert Gottlieb, Pere Alberch, Stephen Jay Gould, Massimo Pigliucci, Robert Lickliter and Mary Jane West-Eberhard — have expressed similar misgivings about this orthodoxy. These individuals, with their deep appreciation of development, have seen the need to expand our evolutionary vision.

Contrary to Coyne's assertions, I never argue that genetics has a minor role in evolution, nor do I suggest that the evolutionary embrace of Gregor Mendel was misguided. On the contrary, in my book I repeatedly discuss the role of genes in normal and anomalous development and the capacity of genetic mutations to produce oddities.

But genes are only part of the answer. And so my argument in *Freaks of Nature* is that we need a more balanced approach. Throughout the book, I invoke the concept of interchangeability to explain, for example, how sex chromosomes or incubation temperature (or even both) can trigger in various species the developmental cascade of events that produce male and female.

In other words, development is a process comprising genetic and non-genetic factors, and evolution has shaped the entire system — not just the genes — to produce the diversity of life forms we see around us. From this perspective, development must have a crucial role in mediating the transmission of form and behaviour from one generation to the next.

Mark S. Blumberg Department of Psychology, University of Iowa, Iowa City, Iowa 52242, USA
e-mail: mark-blumberg@uiowa.edu

How objective is a definition in the subspecies debate?

SIR — Michael Patten, in his Correspondence stating that 'subspecies' and 'race' should not be used as synonyms (*Nature* **457,** 147; 2009), claims that subspecies "remains a useful taxonomic division that enriches our understanding of evolution and biogeography". But, as a classificatory unit, subspecies are not useful in comparative systematic and biogeographical studies because — unlike genera and families, for example — subspecies are groups of populations that are defined by hypothesized biological interactions or geographical distributions, rather than by homology (shared derived characteristics).

Patten says that his definition of subspecies is not arbitrary, "as there are clear ways of describing a subspecies objectively". Yet there are no objective ways to describe species, let alone subspecies.

The species-concept debate is a result of many claiming to have found the 'objective' way to describe a species. So far, this has led to more than 25 species concepts. Patten's definition may therefore represent another addition to the already growing number of 'subspecies concepts'.

Malte C. Ebach International Institute for Species Exploration and School of Life Sciences, College of Liberal Arts and Sciences, Arizona State University, PO Box 874501, Tempe, Arizona 85287-4501, USA
e-mail: mcebach@asu.edu
David M. Williams Department of Botany, The Natural History Museum, Cromwell Road, London SW7 5BD, UK
e-mail: d.m.williams@nhm.ac.uk

Batteries versus biomass as a transport solution

SIR — In his Correspondence 'Choosing between batteries or biomass to stay on the road' (*Nature* **457,** 257; 2009), Lucien Trueb presents figures that do rough justice to the potential of the battery-powered car.

The fully electric car is five times more efficient than the hydrocarbon-guzzling alternative (D. J. C. MacKay *Sustainable Energy — Without the Hot Air* UIT Cambridge, 2008), which takes some of the sting from batteries' lower energy density. And the energy density Trueb quotes for petrol makes no allowance for the mass of the internal-combustion engine: replacing this half-tonne lump with its weight in batteries allows electrification without sacrificing performance. For example, a range of 350 kilometres and acceleration of 0–96 kilometres per hour in less than 4 seconds is obtainable today in commercially available all-electric vehicles for which the price tag, although high, is affordable.

If biofuels are to be the alternative, we will have to sacrifice more than 10% of the world's agricultural land to powering the world's cars, or greatly increase our production of suitable 'waste'. Even then, using bio-hydrocarbons to generate electricity in power stations is roughly twice as efficient as burning them in combustion engines.

The electrification of transport is crucial if we are to break our addiction to fossil fuels. Unlike biofuels, electricity production can be scaled up to keep pace with increased car usage. Electric cars are the only long-term solution to comfortable personal transportation.

Martin Smith Department of Palaeontology, Royal Ontario Museum, 100 Queen's Park, Toronto, Ontario M5S 2C6, Canada
e-mail: martins@rom.on.ca

COMMENTARY

Should scientists study race and IQ?

In the first of two opposing commentaries, **Steven Rose** argues that studies investigating possible links between race, gender and intelligence do no good. In the second, **Stephen Ceci** and **Wendy M. Williams** argue that such research is both morally defensible and important for the pursuit of truth.

P.-A. PETTERSSON/GETTY IMAGES

NO: Science and society do not benefit

Darwin200

Are there some areas of potential knowledge that scientists should not seek out? Or, if they do, should they keep the knowledge secret, hidden from the hoi polloi? Certainly Francis Bacon, that great theorist of the birth of modern science, thought so. For with knowledge comes power — potentially dangerous power. In his utopian novel *The New Atlantis*, scholars determined which of their findings were too dangerous to be shared. Modern governments, obsessed with biosecurity, make similar decisions about what can be researched, how, and in what way disseminated. Private companies bind researchers with non-disclosure and confidentiality agreements. Genetic tests for disorders that have no treatment, such as late-stage Alzheimer's, are often not offered for ethical reasons. As Steven Shapin's book *The Scientific Life* documents, the idea of free, untrammelled and publicly-disseminated research, if it ever corresponded to reality, looks distinctly unrealistic today.

To meet the canons of scientific enquiry a research project must meet two criteria: first, are the questions that it asks well-founded? Research based on the assumption that burning coal releases phlogiston fails this test. And second, are they answerable with the theoretical and technical tools available? As the eminent immunologist Peter Medawar pointed out, science is the art of the soluble. Further, given that our society already accepts a number of restrictions to the pursuit of knowledge, it is sensible to require that funded research also addresses questions that either contribute to basic scientific understanding, offer new beneficial technological prospects, or aid sound public policy-making. These criteria are, of course, those used by both public and private funding bodies.

So what should we make of the century-old but regularly-recycled call for research aimed at discovering whether there are group differences in intelligence?

These days the 'groups' under consideration are 'race' and 'gender'. But it has not always been so. A hundred and fifty years ago, when Darwin published *The Descent of Man, and Selection in Relation to Sex*, he regarded it as so self-evident that white Anglo-Saxon upper-class males were the most intelligent as not to need evidence. Half a century ago, at least in Britain, class was the more relevant grouping, leading to eugenic concerns that the genetically inferior workers were outbreeding their superiors. The issue of race and intelligence became prominent in the United States in the late 1960s, perhaps in response to the civil-rights movement. Arthur Jensen's How Much Can We Boost IQ and Scholastic Achievement? (A. R. Jensen *Harvard Educ. Rev.* **39**, 1–123; 1969) argued that the deficit in black IQ was too great to be explained by deprivation and must be genetic. Similarly, the sex/gender question, naturalized through most of western scientific history, was thrust into the public domain as part of a backlash against emergent feminism in the 1970s by publications such as *The Inevitability of Patriarchy* by Steven Goldberg, which argued that men, by grace of their physiology, were 'naturally' more successful than women at whatever society judged to represent success.

The categories judged relevant to the study of group differences are clearly unstable, dependent on social, cultural and political context. No one, to my knowledge, is arguing for research on group differences in intelligence between north and south Welsh (although there are well-established average genetic differences between people living in the two regions). This calls into question the motivation behind looking for such specific group differences in intelligence, sheds doubt on whether such research is well-founded, and begs whether answers could possibly be put to good use. As we shall see, a more thorough

look at the field will prove that it fails all three of my criteria for justifiable science.

There is a difficulty in the first instance of measuring 'intelligence'. For around a century, this has been done with the IQ test, originally developed in France as a way of supplementing teachers' assessments of their pupils. In the hands of later psychometricians, the tests became increasingly reified, and seemingly made more scientific by the development of the term 'g' to encapsulate 'crystallized' or 'general intelligence'.

However, except to a small band of dedicated psychometricians, it seems obvious that to try to capture the many forms of socially expressed intelligent behaviour in a single coefficient — and to rank an entire population in a linear mode, like soldiers on parade lined up by height — excludes most richly intelligent human activities. Social intelligence, emotional intelligence, the intelligent hands of the craftsman or the intelligent intuition of the scientist all elude the 'g' straightjacket.

"It's just ideology masquerading as science."

The flexibility of IQ

Group comparisons of IQ are even more problematic. Attempts have been made to make 'culture-fair' or 'culture-free' tests, as if such a thing were possible, to allow comparisons of 'g' between people from very different societies. But IQ is clearly a flexible construct — as amply demonstrated by decisions in the 1930s and 1940s in the United States and Britain to 'adjust' test questions to equalize the scores of boys and girls, because in previous versions of the tests girls had scored higher. When Lev Vygotsky tested Russian peasants back in the 1930s, he found that answers that seemed logical to an urbanite were responded to quite differently, but with parallel logic, by the peasants.

As for 'race', the problem is whether it is a biologically, as opposed to socially, meaningful category. Among geneticists interested in differences in gene frequencies between populations, there is increasing consensus that the word obscures more than it reveals, and should be replaced by the concept of biogeographic ancestry, which makes possible the study of subpopulations for relevant genetic and phenotypic characteristics. There are some well-recognized, meaningful genetic differences between groups, for instance between Ashkenazi and Sephardi Jews in terms of their risk to Tay–Sachs disease, and the study of such differences may reveal important clues with respect, for instance, to disease propensity. But such groups are not normally considered socially distinct races for the purposes of studies of group differences in intelligence. Broad divisions between 'white' or 'Caucasian' and 'black' or 'Asian', the groups generally discussed in the context of the IQ debate, especially in the United States, hide genetically important subpopulation differences within these groups.

This is partly why James Watson's ill-advised remark that he was "inherently gloomy about the prospect of Africa" because "all our social policies are based on the fact that their intelligence is the same as ours — whereas all the testing says not really" was recognized not merely as inflammatory, but as scientifically unacceptable in terms of its lumping together of all Africa.

'Gender', that inherently biosocial construct, presents even greater difficulties. In the context of the present debate, the crucial question is whether it is possible to identify a biological — presumably genetic or neurodevelopmental — cause to any difference in the way men and women think and act. The problem is that from the moment of birth, boys and girls are treated differently, which shapes both their growing bodies and brains and how they are expected to behave. It is not just that the biological is expressed through the social and cultural, but that the social and cultural in turn shape the biological. One only has to note how the understanding of what it is to be a man or a woman in Europe or North America has shifted over the past century, to say nothing of how gender relations vary in other cultures. Thus, although there are minor average structural and biochemical variations between Western men's and women's brains (such as the volume of some nuclei and the distribution of hormone receptors), speculations on their implications for how men and women may think or behave lack any empirical basis.

To conclude: the categories of intelligence, race and gender are not definable within the framework required for natural scientific research, failing my first criterion of being

The debate

In this Commentary package, two parties tackle the sensitive proposition that gender- or race-linked differences in intelligence ought to be studied.

Steven Rose, arguing against, is a neuroscientist and biologist, and a founding member of the British Society for Social Responsibility in Science. His books include *Alas, Poor Darwin: Arguments against Evolutionary Psychology* (2000), co-edited with the sociologist Hilary Rose. Stephen Ceci and Wendy M. Williams argue for the proposition. Ceci, who studies the development of intelligence and memory, and Williams, who studies educational and societal implications of intelligence, are co-authors of *The Mathematics of Sex: How Biology and Society Conspire to Limit Talented Women and Girls* (2009).

Neither party saw the other's argument before publication. They will have opportunity to respond to each other and to continue the debate online at http://tinyurl.com/askwhp, where we invite contributions from our readers. **Nature**

PETRIFIED COLLECTION/GETTY IMAGES

BETTMANN/CORBIS

Social and cultural influences have a huge impact on our definitions and measures of race, gender and intelligence.

well-founded. They also fail the second criterion of being answerable: we lack the theoretical or technical tools to study them.

The standard approach of population biologists to estimating the potential genetic contribution to a trait is to make a heritability estimate. Whatever the strengths and weaknesses of this measure within a population, it is essentially just that: a within-population measure, only valid for a given environment. The nature of the equations means that if the environment changes, the heritability estimate changes too. Moreover, the measure relates to a randomly interbreeding population — useful for agricultural purposes such as estimating optimal genotypes for crop yields or milk production — but not for people. Even if reliable correlations were found between some intelligence test score and a measure of brain physiology or activity held by a specific group, such a correlation says nothing about the direction of causation.

As for the third and final criterion, if attempts to answer these group-difference questions are fraught with scientific fallacies, might there nonetheless be some public-policy implications making investigation worthwhile? The answer sometimes advanced is that if there were such differences, and their causes were understood, the less well-endowed groups could be 'compensated' by some form of differentiated education. But in practice, claims that there are differences in intelligence between blacks and whites, or men and women, have always been used to justify a social hierarchy in which white males continue to occupy the premier positions (whether in the economy in general or natural science in particular). Using pseudoscience, based on concepts as ill-founded as was phlogiston, to justify preordained conclusions should not serve as the basis of sound policy-making.

In a society in which racism and sexism were absent, the questions of whether whites or men are more or less intelligent than blacks or women would not merely be meaningless — they would not even be asked. The problem is not that knowledge of such group intelligence differences is too dangerous, but rather that there is no valid knowledge to be found in this area at all. It's just ideology masquerading as science. ■

Steven Rose is a neuroscientist and emeritus professor at the Open University, UK.
e-mail: S.P.R.Rose@open.ac.uk

YES: The scientific truth must be pursued

The Soviet Union lost a generation of genetics research to the politicization of science when Trofim Lysenko, director of biology under Joseph Stalin, parlayed his rejection of Mendelian genetics into a powerful political scientific movement. By the late 1920s, Lysenko had denounced academics embracing Mendelian genetics, which some said undermined tenets of Soviet society. His efforts to extinguish 'harmful' scientific ideas ruined opponents' careers and delayed scientific progress.

It is difficult to imagine this situation repeating today, when rival views feed the scientific process, and inquiry and debate trump orthodoxy. Yet the spectre of Lysenkoism lurks in current scientific discourse on gender, race and intelligence. Claims that sex- or race-based IQ gaps are partly genetic can offend entire groups, who feel that such work feeds hatred and discrimination. Pressure from professional organizations and university administrators can result in boycotting such research, and even in ending scientific careers.

But hatred and discrimination do not result from allowing scientists to publish their findings, nor does censuring it stamp out hatred. Pernicious folk-theories of racial and gender inferiority predated scientific studies claiming genetic bases of racial differences in intelligence. Even if one does not support such work in the interests of free speech, it should be seen as supporting the scientific process of debate. Important scientific progress on these topics would never have been made without the incentive of disproving one's critics.

There is an emerging consensus about racial and gender equality in genetic determinants of intelligence; most researchers, including ourselves, agree that genes do not explain between-group differences. But some issues remain unresolved, such as identification of mechanisms that bring genetic potential to fruition. Censuring debaters favouring genetic explanations of intelligence differences is not the answer to solving such mysteries.

There is a long history to both the study and the vilification of group differences in intelligence. In the late 1800s, Francis Galton, the father of eugenics, was acclaimed and was later knighted for his work. Cesare Lombroso, the scientist who claimed that criminality was inherited and evident in physical features, was also respected during his lifetime. Both were posthumously reviled by some, when their thinking became associated with Hitler's policies, mandatory sterilization and restrictive immigration policies.

"The dominant side goes unchallenged, forestalling the evolution of crucial ideas."

Nobel prizewinner William Shockley became a subject of controversy in the 1970s, after his work turned to racial differences in intelligence. In recent decades, the writings, statements and teachings of Arthur Jensen, Michael Levin and John Philippe Rushton, also on racial differences in intelligence, have met variously with acclaim, outcries and demands for job termination. So have writings of Richard Herrnstein and Charles Murray on the differential distribution of IQ by race. And Frank Ellis, a lecturer at the University of Leeds, UK, took early retirement in the face of an ethical storm that developed after he suggested in a student newspaper that intelligence levels were related to ethnicity. The list goes on. Many have been dissuaded from even looking at the research topic for fear of condemnation.

The outcries against those who speak of racial and gender gaps in IQ have become deafening, at times resembling Lysenkoism in language if not in deed.

Judged too fast?

Consider two recent high-profile cases. In 2005, Harvard's then-president Lawrence Summers suggested gender differences in intrinsic ability as one cause of the dearth of women in the top tier of science, rather than espousing the popular view that women's under-representation results from biased hiring, discriminatory tenure practices and negative stereotypes. Summers's insinuation of biologically-based sex differences in cognitive ability was radioactive, setting off debates on campuses and outpourings of editorials. Despite apologizing for reckless language — which his supporters felt research supported — he later resigned.

James Watson is the most illustrious scholar to have his career ended for reckless language. Watson's downfall was his assertion that "all our social policies are based on the fact that [African] intelligence is the same as ours — whereas all the testing says not really". Although he hoped everybody was equal, "people who have to deal with black employees find this is not true". Watson instantly plunged from A-list Nobelist to outcast, and was suspended from his chancellorship

of Cold Spring Harbor Laboratory. Watson later clarified in a statement that he does not believe Africans to be genetically inferior, but this had little impact on the controversy.

Watson's first assertion could be read as scientifically supported: black Africans' IQ scores are lower than those of white Europeans. But Watson's use of 'intelligence' was interpreted as meaning 'intrinsic cognitive ability', ignoring how unfamiliarity with testing format, low quality of schooling, or poor health might depress IQ scores. There have been analyses showing average national IQs for sub-Saharan Africa to be approximately 30 points lower than average IQs for predominantly white European nations, and drawing a racial conclusion from those results[1,2]. A refutation of these analyses would provide an opportunity to advance understanding. Sadly, although these analyses can be refuted, as we and others have done[3], most of those who scorned Watson never knew they existed.

Attacks on Watson and Summers extinguished discussion by making moral attributions about their presumed character flaws rather than debating facts. But character attacks lead to a one-party science that squelches divergent views.

Some scientists hold more 'acceptable' views, ourselves included. We think racial and gender differences in IQ are not innate but instead reflect environmental challenges. Although we endorse this view, plenty of scholars remain unpersuaded. Whereas our 'politically correct' work garners us praise, speaking invitations and book contracts, challengers are demeaned, ostracized and occasionally threatened with tenure revocation.

Acts of censure edge close to Lysenkoism. They also do a disservice to science. When dissenters' positions are prevented exposure in high-impact journals and excluded from conferences, the dominant side goes unchallenged, and eventually its rationale is forgotten, forestalling the evolution of crucial ideas.

James Flynn, the foremost proponent of the environmental basis of intelligence, notes that when he first rebutted Jensen's hereditarian claims 30 years ago, he never anticipated later breakthroughs that evolved from the debate. Without Jensen, he has written, "I would never have made any contribution to psychology". His landmark documentation of the steady rise in IQ scores across generations and nations, known as the Flynn effect, might never have been done.

Such work has advanced our understanding of intelligence immensely. Flynn showed in 2007 that the IQ of African Americans today is 10–15 points below that of white Americans, but equivalent to the score of such whites in 1947 (the racial groups described here are based on social constructs and rely on study participants defining their own race). Children in the United States and many European nations have surpassed their grandparents' IQs by more than 15 points in 60 years, so it is clearly possible to close the racial gap. This has been happening: 25% of this difference has been eliminated over 30 years[4]. IQ scores of offspring of German women and Second World War black US soldiers have been shown to be indistinguishable from scores of offspring of German women and white US soldiers[4]. No longer are there claims of a linear relationship between IQ and European genes. It is now recognized that cultural effects are more powerful than previously thought.

BETTMANN/CORBIS

IQ tests got a bad name through association with immigration policy.

Regarding gender, no one now claims women are unable to excel at complex maths: 48% of US mathematics majors are female, and women earn higher maths grades than men throughout schooling[5]. The maths gender gap among the top 0.01% of students, which 30 years ago favoured males 13-to-1, now favours males only 2.8-to-1 (ref. 5). Some nation's women (including those in Singapore and Japan) outscore US males on maths tests by an amount far larger than the gender gap within the United States[5].

So, vigorous debate has resulted in great progress in our understanding, and more breakthroughs will come — if we allow free speech in science.

One could argue that some peer-reviewed reports feed racial and gender stereotypes. Perhaps such research should be forced to pass a higher cost–benefit threshold before publication. But this is a slippery slope: philosopher Jean-Jacques Rousseau argued that atheists should be silenced, lest they convince the masses to abandon faith, condemning them to hell. This would now be viewed as a ludicrous violation of free speech. Who is to be impanelled with the wisdom to decide which views can be aired, and which research questions pass muster?

It might also be argued that only primary researchers who are experts in their field, rather than administrators or non-experts, deserve protected speech in these areas. A statement's validity, however, lies in its congruence with scientific data, not in the role occupied by its speaker.

One powerful argument states that groups need protection against bigotry, and that censuring one side in a debate is necessary to prevent the harm done to victims of race and gender arguments. Certainly, history offers examples of great harm befalling individuals due to flawed scientific claims. Such problems, however, arise not from scientific discourse, but from political applications of those ideas. This is another matter entirely and must be subject to checks and balances.

In today's world, subjective perceptions of scientists' intent seem to determine a study's acceptability — work is celebrated if perceived as elevating under-represented groups (as with focuses on women and minorities in the search for personalized medicine), but reviled if perceived as documenting sex and race differences in intelligence without a focus on interventions to eliminate them. Yet many future uses of knowledge cannot be anticipated; Flynn's research has since been used to overturn death-row sentences for mentally-retarded, disproportionately black defendants, for example.

When scientists are silenced by colleagues, administrators, editors and funders who think that simply asking certain questions is inappropriate, the process begins to resemble religion rather than science. Under such a regime, we risk losing a generation of desperately needed research. ■

Stephen Ceci and **Wendy M. Williams** are in the Department of Human Development at Cornell University, Ithaca, New York, USA.
e-mails: sjc9@cornell.edu; wmw5@cornell.edu

1. Lynn, R. & Vanhanen, T. *IQ and the Wealth of Nations* (Praeger, 2002).
2. Lynn, R. & Vanhanen, T. *IQ and Global Inequality* (Washington Summit, 2006).
3. Barnett, S. M. & Williams, W. M. *Contemp. Psychol.* **49,** 389–396 (2004).
4. Flynn, J. R. *What is Intelligence?* (Cambridge Univ. Press, 2007).
5. Ceci, S. J., Williams, W. M. & Barnett, S. M. *Psychol. Bull.* **135,** 218–261 (2009).

See Editorial, page 763, and online at www.nature.com/darwin.

ESSAY

A flight of fancy

Darwin200

Henry Nicholls wonders how things would be different had Charles Darwin given in to pressure from his publisher to rewrite *Origin of Species* into a popular book about pigeons.

"Pigeons indeed," huffed Charles Darwin, his brow furrowed as he read to the end of a letter and laid it down on his desk.

The letter contained feedback from his publisher John Murray on a draft of what would become *On the Origin of Species by Means of Natural Selection*. Murray had farmed out copies of the manuscript to a couple of his trusted advisers. One of them, a rural vicar and literary editor by the name of Whitwell Elwin, had not liked it all. "At every page, I was tantalized by the absence of the proofs," Elwin had written to Murray on 3 May 1859. In contrast to the *Journal of Researches* (later known as *Voyage of the Beagle*), which Elwin had found "one of the most charming books", Darwin had written this new work in a "much harder & drier style".

Although opposed to the publication of what he saw as "a wild and foolish piece of imagination", Elwin hadn't advised Murray to reject the manuscript outright. Instead, he had sought the advice of the geologist Charles Lyell. It was Lyell who said that the book should focus on Darwin's observations of pigeons "accompanied with a brief statement of his general principles" on natural selection. Indeed, Lyell had made a similar suggestion to Darwin back in May 1856 after being enchanted by Darwin's pigeons at Down House, Kent. "I wish you would publish some small fragment of your data [on] pigeons if you please & so out with the theory & let it take date — & be cited — & understood," he had written.

Elwin's final recommendation to Murray reiterated this wish: Darwin should deliver a slimmer volume that homed in on this one well-worked case study. "Every body is interested in pigeons," Elwin observed.

Before considering Darwin's response to that suggestion, it is worth taking a moment to reflect on what might have been *On the Origin of Pigeons*. Such a volume would, Elwin had suggested, "be reviewed in every journal in the kingdom & would soon be on every table". He was probably right. Breeding 'fancy' pigeons was an extraordinarily popular pastime in Victorian Britain, with enthusiasts spanning the entire social spectrum, from the poorest weavers in London's Spitalfields to Queen Victoria herself. But how effective would *On the Origin of Pigeons* have been as a vehicle for Darwin's ideas on evolution by natural selection?

Despite their striking physical differences, Darwin suspected that the various breeds of 'fancy' pigeon shared a common ancestor.

PIGEONS: D. WOLSENHOLME/DOWN HOUSE, KENT, UK/THE BRIDGEMAN ART LIBRARY

Common descent

It was just four years before he received Murray's letter, in March 1855, that Darwin found himself drawn towards fancy pigeons. For the pigeon breeders, or fanciers, crossing different breeds was not just pointless (as mixed-race offspring lack the qualities of their parents), it bordered on heresy. So like the splendid isolation offered by the different islands of the Galápagos, the strict racial segregation the fanciers imposed on their birds ensured that every breed was effectively isolated from every other.

Despite the birds' striking physical differences, Darwin, in line with several other naturalists, suspected a single common ancestor for all the varieties — the rock pigeon *Columba livia* — and he began to gather evidence to support his hunch. Within months, he had mugged up on manuals, tailor-made a pigeon house for his garden and filled it with dozens of different breeds, describing them to Lyell as "the greatest treat, in my opinion, which can be offered to [a] human being". With abundant historical information at his fingertips, and by obtaining meticulous measurements from adults, embryos and even blood corpuscles, he constructed a family tree for the pouters, carriers, runts, barbs, fantails, nuns and a dozen other spectacular fancy pigeon breeds.

Darwin's foray into breeding gave him yet more evidence in favour of the pigeons' common descent. A cross of any two varieties, he found, produced fertile offspring. What's more, the offspring of these crosses often sported characteristics not found in either parent but that did appear in the rock pigeon. So in spite of their extraordinary visible differences, some underlying sameness united all the fancy varieties. All this, and more, eventually wound up in his 1868 book on *The Variation of Animals and Plants Under Domestication*.

Darwin's pigeons also served another

purpose, one that's harder to detect in his published writings. By crossing different breeds and observing how traits were passed from one generation to the next, he hoped to bring some clarity to the frequently confused thinking on inheritance. Research on the mass of pigeon skeletons and skins that Darwin left to the Natural History Museum in London promises to shed fresh light on the precise lines of enquiry he pursued. But even with more information, seeing pigeons as Darwin saw them will be extraordinarily tricky because it requires unimagining so much of what we now take for granted. At the very least, you must unravel the double-helix, knock out the gene and banish the concept of chromosomes. Then you must populate your mind with what are, by today's standards, some very strange notions.

Take pre-potency, for example, the idea that when two different breeds are crossed, one makes a greater contribution than the other to the offspring. In a version of this, known as Yarrell's law, it was proposed that the older breed would always dictate the characteristics of the next generation. Another commonly held belief was that males were responsible for defining the external characteristics of the offspring and females the internal features.

Evidence from Darwin's notebooks, letters and unpublished manuscripts suggests that the outcomes of the crosses he performed ran counter to each of these widely held beliefs. Unfortunately, however, the complex genetic basis for the traits he was looking at meant his pigeon crosses failed to reveal the true particulate nature of inheritance, something that would have to wait for the discovery of Gregor Mendel's work at the turn of the century.

Publisher John Murray (top) relayed to Darwin the advice of Charles Lyell, who felt that pigeons would best communicate his theory of natural selection.

Unnatural selection

Although Darwin's work on pigeons failed to unravel many of the mysteries surrounding inheritance, it did make a marvellously persuasive case for the origin of new varieties by means of artificial selection. From there, it was a simple step to argue, by analogy, for the origin of species by means of natural selection. But would *On the Origin of Pigeons*, had it been written, have conveyed the enormity of Darwin's idea?

> **"On the Origin of Pigeons would have failed to capture the extraordinary scope of natural selection."**

A few of the more observant and imaginative readers might well have seen Darwin's broader message. For many more, however, the wonder of the fancy breeds would have been the main show, and the link between artificial and natural selection an unnecessary distraction. Perhaps more importantly, *On the Origin of Pigeons* would have failed to capture the extraordinary scope of natural selection's explanatory power.

That was Darwin's crucial aim — to demonstrate the impact of natural selection "on geograph. distribution, palaeontology, classification Hybridism, domestic animals & plants…". In 1856, he set about writing the book that would do this, a work he called *Natural Selection*. He was still at it almost two years later when he received the now infamous letter from Alfred Russel Wallace that outlined the same idea. This unexpected circumstance forced Darwin to shelve what would have been a massive, multi-volume magnum opus (about two-thirds of which he had already written), and work furiously for the next ten months to bash out a shorter 'abstract'.

So how did Darwin feel that morning when he received Elwin's feedback on this condensed version? Probably a bit rotten. Elwin's criticisms — that it was thin on evidence and stylistically wanting — hit on the two weaknesses that he often worried about. But if he did contemplate writing the desired pigeon book, it was not for long. "It is my deliberate conviction that both Lyells & Mr Elwyns suggestions … are impracticable," he wrote back to Murray. "I have done my best. Others might, I have no doubt, done the job better, if they had my materials; but that is no help."

Persuaded by Darwin's determination, Murray agreed to publish, making little editorial input other than to clip the first five words from the proposed title — *An Abstract of an Essay on the Origin of Species and Varieties Through Natural Selection*. Darwin subsequently made a few further tweaks to give us the title we celebrate this year.

Darwin's gambit — to convince his audience of evolution by natural selection by showing the extraordinary breadth of its explanatory power — proved remarkably effective. We owe a debt of gratitude to Wallace for denying us the thudding great tomes of *Natural Selection* and to Darwin for depriving us of a pithy little book on pigeons. *On the Origin*, with its light touch and wide-ranging content, sits comfortably between these two extremes. ■

Henry Nicholls is a science writer who lives in London. His most recent book is *Lonesome George: The Life and Loves of a Conservation Icon*.

For further reading, see http://tinyurl.com/bwjsl5. For more on Darwin, see www.nature.com/darwin. Listen to Henry Nicholls on the *Nature* podcast at www.nature.com/nature/podcast.

BOOKS & ARTS

Darwin and his family would have celebrated the end of colonial slavery, depicted in this 1834 illustration.

BODLEIAN LIBRARY OXFORD/THE ART ARCHIVE

A vision of humanity united

A controversial new reconstruction of Charles Darwin's life suggests his family's campaign against slavery influenced his belief that all humans evolved from a single ancestor, explains **W. F. Bynum**.

Darwin200

Lewis Carroll wrote of a mythical island whose inhabitants made a precarious living by taking in each other's laundry. Most books on Charles Darwin in his bicentennial year will probably follow suit, repackaging what a troop of Darwin scholars have uncovered from his extensive archives and writings. The more books that have been written on a subject, the easier it seems to be to write a new one. But *Darwin's Sacred Cause* is an exception.

Readers of Adrian Desmond and James Moore's earlier biography, *Darwin* (Michael Joseph, 1991) — and of Desmond's subsequent *Huxley: From Devil's Disciple to Evolution's High Priest* (Penguin, 1998) — will expect much from this new reading of Darwin's life and values. They will not be disappointed. In *Darwin's Sacred Cause*, Desmond and Moore assimilate the relevant secondary literature, but also go much further. They offer us a new reading of Darwin's life and scientific work, based on two well-known facts about him: he felt physical revulsion when confronted by cruelty and he loathed slavery.

Darwin's hasty exit from the surgical theatre, while a student in Edinburgh, is a classic anecdote of his early years. Even necessary pain, such as that caused by operations before anaesthetics were available, was more than he could bear. And no one can read his *Journal of Researches*, later known as *Voyage of the Beagle*, without realizing how much human slavery affected his sensitivities. He feared that his book might be censored in the United States, where slavery was a major political and economic issue.

Darwin's Sacred Cause: Race, Slavery and the Quest for Human Origins
by Adrian Desmond and James Moore
Allen Lane/Houghton Mifflin: 2009.
512 pp/448 pp. £25/$30

Desmond and Moore have reconstructed Darwin's life and science through the lenses of these two themes. This approach has many pay-offs. First, it enlarges our knowledge of Darwin's circle. He came from a family with ardent anti-slavery sentiments. Darwin's grandfather, Josiah Wedgwood, produced the most powerful image of the British anti-slavery campaign:

a jasperware medallion of 1787, depicting a slave in chains, which bore the motto 'Am I not a Man and a Brother?' (pictured, right). Darwin's sisters, aunts, uncles and cousins all wholeheartedly supported the 'sacred cause' with their energies and their money. The authors mine correspondence, pamphlets and monetary contributions to dissect this family concern in a detail never before achieved, and demonstrate convincingly that Darwin absorbed these family values.

This concentration on Darwin's abiding belief in the unity of the human races, in the wake of a rising tide of polygenism — the belief that humans were divided into different races with separate origins — encourages us to rethink his relationships with his acquaintances and colleagues. The doughty Harriet Martineau, his brother Erasmus's friend, becomes not just a source of Malthusian commentary, but an ardent critic of the iniquities of American slavery, based on first-hand experience during a long visit to the United States. Darwin's complex association with Charles Lyell is shown to have very deep and complicated roots. We know that Darwin was extremely disappointed with Lyell's reception of *On the Origin of Species*. Desmond and Moore dissect correspondence during Lyell's first two trips to North America, where the patrician geologist was wined and dined by aristocratic slaveholders in the American South. Their analysis of these years in the 1840s is both amusing and moving, as Darwin goaded Lyell to see beyond the polite trappings of southern society to the harsh realities on which it was based.

A personal evolution

Desmond and Moore also invite us to reconsider Darwin's attitude to the naturalist Louis Agassiz. His rejection of Darwin's central message in *On the Origin of Species* is well known, but we can now see the deep roots of Darwin's antipathy to this Harvard apostle of polygenism, who wrote an introduction to the most prominent racist tract in antebellum America, *Types of Mankind* by J. C. Nott and G. R. Gliddon (1854). According to Desmond and Moore, Darwin canvassed first the zoologist James Dana at Yale, and then, when Dana proved too much in Agassiz's shadow, the botanist Asa Gray at Harvard, to act as his undercover agent in Agassiz's America. They argue that Agassiz's polygenism, rather than his espousal of continental, idealist modes of thinking, was what Darwin found especially odious.

This account of Darwin's personal evolution also alters the focus on the important influences on his thinking. In particular, the Bristol physician and ethnologist James Cowles Prichard (1786–1848) becomes a major player. Darwin certainly owned and read carefully many of Prichard's books, especially his *Researches into the Physical History of Mankind*, a monumental work of erudition that was transmuted from his University of Edinburgh MD thesis (1808) to a single volume (1813), and finally, in its last incarnation, into a five-volume encyclopaedic compendium of biology, linguistics, ethnology and physical anthropology. Prichard was a Quaker-turned-Evangelical Anglican, and certainly no evolutionist. Nevertheless, he was an ardent believer in the unity of mankind, marshalling all the evidence he could that all human races were derived from a single stock. Prichard invoked what he called the 'analogy of nature', by which he meant that humans, similar to domestic animals (dogs, sheep, cattle), vary far more than species in the wild. Mankind, like the dog, is domesticated, and therefore capable of wide variation.

WILBERFORCE HOUSE, HULL CITY MUSEUMS AND ART GALLERIES, UK/BRIDGEMAN ART LIBRARY

Given the importance Darwin placed on domestication as a model of biological variation, the symmetry of their arguments is stark. Darwin noted at the back of one of his volumes of Prichard, 'How like my Book all this will be'. The comment is famous and has been variously interpreted by historians. Desmond and Moore argue that Darwin intended his treatment of human variation, albeit within an evolutionary context, to read much like Prichard's. The extensive documentation contained in their book makes this interpretation plausible. Certainly Darwin found in his ethnological reading much appeal to the 'analogy of nature' among those wanting to account for human variation as being on a par with the variation still produced among domestic animals.

Prichard argued that skin colour varies with the degree of 'civilization', and proposed that Adam was black. He also had a notion of sexual selection, whereby ideas of beauty, different in specific cultural contexts, would have reinforced racial variations. Darwin, too, in Desmond and Moore's account, saw sexual selection as the key to varieties of mankind. They suggest that Darwin abandoned a chapter on the races of man in his planned book on natural selection because he felt he needed more information on sexual selection. In the event, the planned book was overtaken by Alfred Russel Wallace's letter revealing that he had come to similar conclusions independently, and the findings of the two men were famously presented together at a meeting of the Linnean Society in 1859. The authors provide good evidence for this, although there is rather more on sexual selection in the first edition of *On the Origin of Species* than the uninitiated reader would conclude from *Darwin's Sacred Cause*.

Back to his roots

The 'analogy of nature' presupposes that all domestic plants and animals — humans as well as pigeons, dogs and sheep — are varieties of a single original species. Prichard assumed it was true, and Darwin strove manfully to follow in his footsteps. Darwin's researches on pigeons were a tour de force, and never written about with more force than here. The case for dogs was more complicated, and his conclusion, in *On the Origin of Species*, that our domestic dogs are descended from more than one species, was a problem — especially given the importance of his notion of descent with variation. On Desmond and Moore's reading of his values, with its commitment to human unity, dogs presented a real difficulty.

The natural culmination of this volume is Darwin's *Descent of Man and Selection in Relation to Sex* (Murray, 1871). Desmond and Moore reinforce their point that Darwin's concern with sexual selection had some of its origins in his abiding belief in the fundamental unity of the human races, but even these authors cannot rehabilitate Darwin's disappointing discussion of human descent. During the 1850s and 1860s, evidence for the antiquity and early history of humans had been established. Too many of Darwin's leading ideas were still grounded in his own amazingly creative period in the late 1830s. After *On the Origin of Species*, much of Darwin's innovative research was in botany; it is as if he wanted most of all simply to be in his own garden. ■

W. F. Bynum is emeritus professor of the history of medicine at University College London, Gower Street, London WC1E 6BT, UK. He is author of *The History of Medicine: A Very Short Introduction* and editor of a new edition of *On the Origin of Species*.
e-mail: w.bynum@ucl.ac.uk

For more on Darwin, see Editorial, page 763, and online at www.nature.com/darwin.



Poems from Darwin's descendant

Amid the many analyses of Darwin's life and work, a more intimate literary portrait emerges from the poetry of his great-great-granddaughter, **Ruth Padel**. Her series of poems on his life — six of which are reproduced here — evokes the emotion and drama of the naturalist's discoveries.

'The Miser'

Cross the Welsh Bridge out of town, go up the hill
on Frankwell Street and you'll see, above the Severn,
brick pillars with the sandy bloom of an ageing dog.
Around the back, Father's surgery and waiting-room.

Outside, the Stable Yard: hay chutes, a piggery and toolshed.
Lower down, a bothy on the river bank
where plates of jagged ice, harvested in winter from the river,
lean one against the other. A dairy where these blocks are dragged

to cool the milk and cream. The Quarry Pool
where he fishes for newts and tadpoles.
Collecting: to assert control
over what's unbearable. To gather and to list.

"Stones, coins, franks, insects, minerals and shells."
Collect yourself: to smother what you feel,
recall to order, summon in one place.
Make a system, like Orpheus, in the face of loss.



'Plankton'

The deck is dazzle, fish-stink, gauze-covered buckets.
Gelatinous ingots, rainbows of wet flinching amethyst
and flubbed, iridescent cream. All this
means he's better; and working on a haul of lumpen light:

polyps, plankton, jellyfish and sea butterflies, the pteropods.
"So low in the scale of nature, so exquisite in their forms!
You wonder at so much beauty — created,
apparently, for such little purpose!" They lower his creel

to blue pores of subtropical ocean. Wave-flicker, white
like a gun-flash over the blown heart of sapphire.
Peacock eyes, beaten and swollen,
tossing on lazuline steel.

'Like Giving to a Blind Man Eyes'

He's standing in Elysium. Palm feathers, a green
dream of fountain against blue sky. Banana fronds,
slack rubber rivulets, a canopy of waterproof tearstain
over his head. Pods and racemes of tamarind.
Follicle, pinnacle; whorl, bole and thorn.

"I expected a good deal. I had read Humboldt
and was afraid of disappointment."
What if he'd stayed at home? "How utterly vain
such fear is, none can tell but those who have seen
what I have today." A small rock off Africa —

and alone with his enchantment. So much and so unknown.
Like taking a new-born baby in your arms. "Not only the grace
of forms and rich new colours: it's the numberless —
& confusing — associations rushing on the mind
that produce the effect." He walks through hot damp air

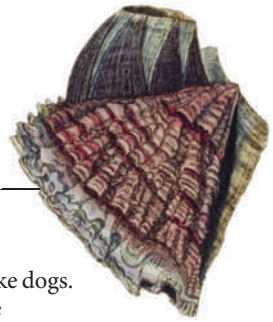
and tastes it like the breath of earth; like blood.
He is possessed by chlorophyll. By the calls of unknown birds.
He wades into sea and scares an octopus. It puffs black hair
at him, turns *red* — as hyacinth — and darts for cover.
He sees it watching. He's discovered

something wonderful! He tests it against coloured card
and the sailors laugh. They know that girly blush!
He feels a fool — but look, he's touched Volcanic rock
for the first time. And Coral on its native stone.
"Often at Edinburgh have I gazed at little pools

of water left by tide. From tiny Corals of our shores
I pictured larger ones. Little did I know how exquisite,
still less expect my hope of seeing them to come true.
Never, in my wildest castles of the air, did I imagine this."
Lava must once have streamed over the sea-floor here,

baking shells to white hard rock. Then a subterranean force
pushed everything up to make an island. His first evidence
of *Volcano*! Vegetation he's never seen, every step a new surprise.
"New insects, fluttering about still newer flowers. It has been
for me a glorious day, like giving to a blind man eyes."

BARNACLES: ENGLISH HERITAGE PHOTO LIBRARY; OTHERS: PRIVATE COLLECTION/BRIDGEMAN ART LIBRARY



'Algae from the Arctic'

On high trails, the mules leave scarlet dints in snow
as if their feet had bled. He scoops up the wet red
and rubs it. It isn't claret dust of porphyry, blown
from peaks of the Peuquenes. Nor granite
from the blue-roan Portillo ridge. No —

none of this heart-sweeping geology, the Cordillera —
ashy wind, guanaco flocks
and level-soaring condors
which have bewitched him on this jaunt
to the interior. (He's done so many, now.)

"I felt glad I was by myself. Bright-coloured rocks,
profound valleys, heaped ruins
and wild broken forms — like watching
a thunderstorm or hearing in full Orchestra
a chorus from *The Messiah*."

In the microscope, ruby specks glow
in ice the mules have squashed. "Like eggs
of small molluscous animals." Or midget ticks
full of blood. With Covington, he posts them off
to Cambridge — to be identified as algae, from the Arctic.

'On Not Thinking About Variation in Tortoise-Shell'

Pure volcano. A mantle of hot bare rock. "Nothing could be less
inviting. A broken field of black basaltic lava
thrown into most rugged waves and crossed
by fissures." Lava tubes, tuff cones and bright,
red-orange crabs. A land iguana! One saffron
leathery elbow, powdery as lichen, sticking out

like a man doing press-ups while leering at the sand.
And the marine iguana … "Hideous! An imp of darkness.
On Albermarle they seem to grow to a larger size."
Young sea lions nip their tails for fun and fling them in the air
like cats with mice. To eat them? No — nothing here,
except one hawk, is carnivore; and none afraid of Man.

Look — giant tortoises! "Travelling eagerly, their necks
outstretched, to springs. I tried riding on their backs
but found it hard to balance! The colony's Vice-Governor
told us he knew which island any shell came from
because they differ. I did not for some time
pay this enough attention. I never dreamed

that islands sixty miles apart, made of the same stone,
of nearly equal height in the same climate,
could have different tenants." Fast forward twenty years
and you see him write of this scatter-burst of rock in open sea,
"We seem brought near that mystery of mysteries,
the first appearance of new beings on the earth."

'Painting the Bees'

The questions buzz at him like birds. They cling
like burrs, delight him like the children, paw like dogs.
They scratch, torment and swarm; they pollinate
like wasps. It's got to be vast — proof, evidence,
minutiae. Orchids, fertilizing in the greenhouse.
Birdskins from India, the horse-markings of Norway, finch
beaks from Galapagos, a parcel of flora from Kew.

How contain all this? Vomit. Panic. Vomit. Nerves
in his neck. That wagtail on the lawn —
what muscles make it flick like that; what purpose
does it serve? Flatulence. Bad headache. Halfway through.
He steps out through the garden to his thinking path,
the Sandwalk. Remembering the trail at Maer
around the lake, they have designed a circuit round a copse.

Chasing each question through the undergrowth
with his fox terrier, he leaves a pebble at the start
of every lap. Into The Dark Side (so the children call it
and the little ones are scared). Ivy, tall trees, bushes,
lords and lady berries, evergreens. Out past
the sandpit to the Light Side's open view: of sun
and cloudshadow over Great Pucklands Meadow.

He loves that children play here. A bugle and boy-laughter —
that's Frank and Georgy in Crimea. He's drilled them with toy rifles,
taught them parts of a gun. Since *Beagle,* he hasn't used one.
("Shooting is cruel.") They're warming their campfire
milk. "Whatever my father did with us," they will say after,
"had over it a glamour of delight." His Fool's Experiments!
"I shan't be easy till I've tried …" — as if some outside force

were at him with a truncheon. He sets the girls to search
for wormcasts. "Damp evenings are best."
Horace finds him snake and lizard eggs, Frank plays bassoon
to worms (are they really deaf?) and even flowers
to see how they like vibration. He strings all seven children over
long grass and scabious in a chain, to paint
and track the bumblebees who pollinate red clover.

CARPENTER-TURNER

Darwin: A Life in Poems
by Ruth Padel
Chatto & Windus: 2009. 160 pp. £12.99

Prizewinning poet Ruth Padel is Charles Darwin's great-great-granddaughter. Stories told to her by her grandmother, who edited several of Darwin's books, inspired her to compose a series of poems for the bicentenary of the naturalist's birth, a selection of which are reproduced here.

Padel's vivid poetry offers an intimate portrait of Darwin through his life — as a young man marvelling at nature in his own backyard; as an explorer exclaiming on seeing a tropical rainforest for the first time; as a researcher puzzling over his collections in his home office at Down House in Kent, being interrupted by his children playing. Her insight exposes Darwin's thoughts as he pieced together his ideas on evolution, and reflects on how his changing views affected his relationships with his family, the church and society.

To hear Ruth Padel reading her poetry, download the *Nature* Podcast at http://tinyurl.com/faze2. For more on Darwin, see www.nature.com/darwin.

NEWS & VIEWS

VIROLOGY

Final entry key for hepatitis C

Thomas Pietschmann

To propagate, the hepatitis C virus relies on entry into liver cells. Ironically, cellular proteins that normally assemble firm seals between adjacent cells serve as crucial keys by which the virus gains access.

Viral replication is possible only if the virus has access to cellular resources. But not all cells allow this. Usually, only certain cell types of a specific host species carry all the necessary cofactors to render them susceptible to infection with a distinct virus. For example, the hepatitis C virus (HCV), which causes severe liver disease, infects only humans and chimpanzees. Within these hosts, it replicates in the main type of liver cell, the hepatocyte. Consequently, until now studies of this virus in animal models have been restricted to chimpanzees and to immunodeficient mice transplanted with human hepatocytes. On page 882 of this issue, Ploss *et al.*[1] identify a host-specific protein that is essential for HCV entry into cells. If this protein could be introduced into small animals by genetic engineering, to produce models of HCV infection, research into fighting this infection would be furthered, as manipulating the protein in humans might be a route to protection against the disease.

A decade ago, HCV entry into cells seemed fairly simple: the virus was thought to bind to a single specific protein on the host-cell surface. So when CD81 was identified[2] as a cellular protein that binds the main protein (E2) on the surface of HCV, it seemed that the key to HCV entry had been discovered. Now, however, the list of factors involved in HCV entry has grown, and includes at least two other essential proteins — SR-BI and claudin-1 — and potentially additional accessory factors such as glucosaminoglycans and low-density-lipoprotein receptors[3–5]. Nonetheless, even when cells are manipulated to express all of these factors together, they can still resist HCV infection[5].

This observation indicated that at least one other host factor crucial for HCV entry was out there. In their search for this missing key, Ploss *et al.*[1] have hit the mark a second time. Following the discovery of claudin-1 by some of the same authors, Ploss and colleagues have now used a similar screen to identify occludin as a host-cell protein essential for HCV entry. Remarkably, both claudin-1 and occludin are components of tight junctions — structures that form firm seals between adjacent cells. Interacting in a zipper-like fashion, these proteins form strands between adjoining cells, sewing them together. By means of these contacts, cellular sheets such as the layer of epithelial cells lining the blood vessels prevent leakage. Moreover, tight junctions form a 'fence' in the cellular membrane, allowing cells to maintain functionally distinct surface-membrane domains. In the case of hepatocytes, these junctions separate a canalicular domain engaged in bile secretion from a luminal domain that is in constant contact with the blood (Fig. 1a).

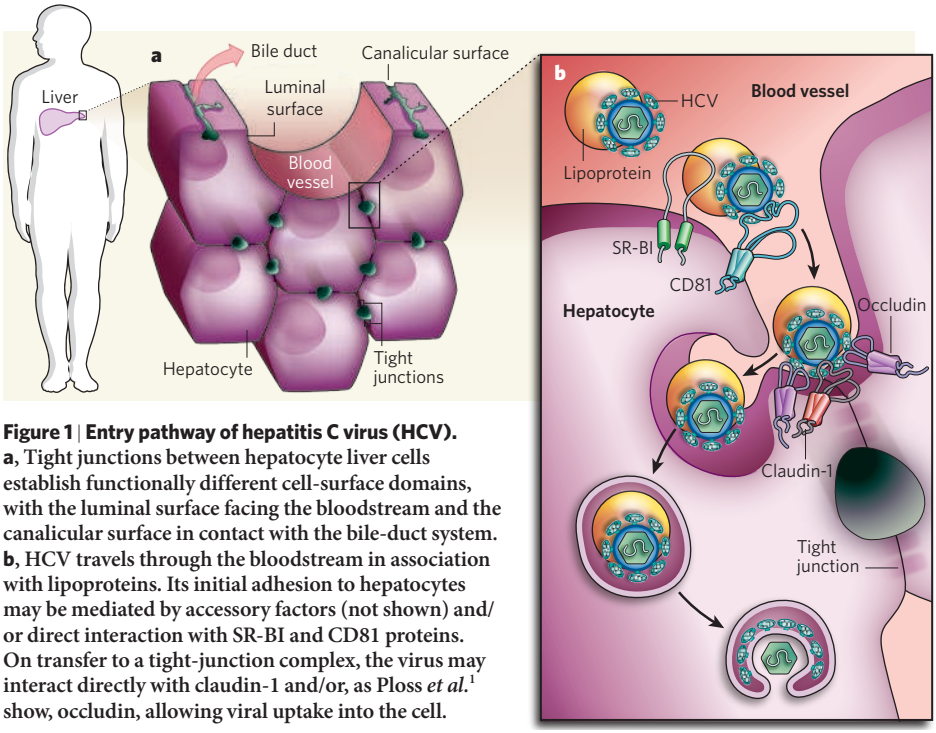


Figure 1 | Entry pathway of hepatitis C virus (HCV). **a**, Tight junctions between hepatocyte liver cells establish functionally different cell-surface domains, with the luminal surface facing the bloodstream and the canalicular surface in contact with the bile-duct system. **b**, HCV travels through the bloodstream in association with lipoproteins. Its initial adhesion to hepatocytes may be mediated by accessory factors (not shown) and/or direct interaction with SR-BI and CD81 proteins. On transfer to a tight-junction complex, the virus may interact directly with claudin-1 and/or, as Ploss *et al.*[1] show, occludin, allowing viral uptake into the cell.

Viral exploitation of occludin and tight junctions for entry into cells is not unprecedented. Coxsackievirus B, for example, first binds to its main receptor at the luminal surface of epithelial cells, and this receptor–virus complex is then transported to tight junctions, where it interacts with a second receptor in an occludin-dependent fashion to initiate the virus's uptake[6,7]. Similarly, HCV may use other proteins for binding and travelling to the tight junctions, where contact with claudin-1 and occludin may provide the final keys for viral entry into the cell (Fig. 1b). The virus might choose this mode of entry because it follows a natural route of cargo delivery into hepatocytes, with the virus having evolved ways to usurp the cellular proteins mediating this process. It will be interesting to learn whether other liver pathogens, including hepatitis B virus and plasmodium sporozoites of the malaria parasite, target this organ using similar routes of invasion. Intriguingly, at least two proteins that usher HCV into hepatocytes (SR-BI and CD81) also participate in the uptake of plasmodium[8–10].

Ploss and colleagues' identification of occludin probably ends the search for essential factors for HCV entry: as long as human CD81, SR-BI, claudin-1 and occludin were present, every cell the authors tested became infected with HCV. This finding indicates that all other cellular proteins that might participate in HCV invasion are ubiquitous. Although none of the four essential proteins for HCV entry is exclusive to liver cells, only a few cells may have all four at their disposal. Presumably, the large number of essential entry factors and their relative

abundance on hepatocytes partially explain HCV's preference for infecting the liver.

CD81 proteins from other mammals, such as the mouse, are inefficiently used by HCV[11]. Such obstacles to the virus's entry limit its host range. Ploss *et al.* show that, although HCV does not discriminate between human and mouse SR-BI and claudin-1, mouse occludin — like its CD81 protein — cannot substitute for the related human protein in aiding viral entry. Does this mean that transgenic mice engineered to produce human CD81 and occludin in their liver cells will host HCV and so can make simple experimental models?

Given the extensive dependence of the virus on its host cell, not only during penetration, but also throughout its replication cycle, overcoming the species-specific entry barrier alone is unlikely to be sufficient for viral propagation. Nevertheless, HCV genomes artificially introduced into mouse cells can replicate, albeit with reduced efficiency[12,13]. Possibly, in the same way that an 'elite' class of factors is required for HCV entry, a small set of host factors exists that is essential for species-specific viral replication and for assembly of viral progeny. Knowing all the essential factors for HCV entry, and that species incompatibility of CD81 and occludin prevent this virus from entering mouse cells, researchers may have reached first base on the path towards a transgenic mouse model for HCV replication. ■

Thomas Pietschmann is at the Twincore Centre for Experimental and Clinical Infection Research, and the Medizinische Hochschule Hannover, D-30625 Hannover, Germany. He is also at the Helmholtz Centre for Infection Research (HZI), Braunschweig, Germany.
e-mail: thomas.pietschmann@twincore.de

1. Ploss, A. *et al. Nature* **457,** 882–886 (2009).
2. Pileri, P. *et al. Science* **282,** 938–941 (1998).
3. Dubuisson, J., Helle, F. & Cocquerel, L. *Cell Microbiol.* **10,** 821–827 (2008).
4. Scarselli, E. *et al. EMBO J.* **21,** 5017–5025 (2002).
5. Evans, M. J. *et al. Nature* **446,** 801–805 (2007).
6. Coyne, C. B. *et al. Cell Host Microbe* **2,** 181–192 (2007).
7. Coyne, C. B. & Bergelson, J. M. *Cell* **124,** 119–131 (2006).
8. Rodrigues, C. D. *et al. Cell Host Microbe* **4,** 271–282 (2008).
9. Yalaoui, S. *et al. Cell Host Microbe* **4,** 283–292 (2008).
10. Silvie, O. *et al. Nature Med.* **9,** 93–96 (2003).
11. Flint, M. *et al. J. Virol.* **80,** 11331–11342 (2006).
12. Uprichard, S. L. *et al. Virol. J.* **3,** 89 (2006).
13. Zhu, Q., Guo, J. T. & Seeger, C. *J. Virol.* **77,** 9204–9210 (2003).

QUANTUM OPTICS

A grip on entanglement

John C. Howell

Quantum buffers will be an essential part of quantum-information networks. A buffer that can preserve not only a 'quantum bit' but also a 'quantum image' is a major step towards creating those networks.

The long-term vision for quantum computation and communication is the use of these technologies as the basis of a quantum-information network. Quantum computers interlinked by quantum communication would form a quantum equivalent of the Internet. With the possibility of unbreakable encryption, simulations of large quantum-mechanical systems and exponentially faster computing for certain algorithms, quantum-information processing offers an important niche in information technologies, and may even be the technology of the future as classical systems reach their physical limits (for example, in transistors, the size of an atom). But, unlike their classical counterparts, all elements in the quantum network require the preservation not only of bits, but also of quantum bits.

A classical bit can have values of 0 or 1, but a quantum bit (qubit) can have a value of 0 or 1, or 0 and 1. Quantum repeaters, buffers, memories and so on all require that a quantum bit is preserved while in that element. An important advance, made by Marino *et al.*[1] and reported on page 859, is the demonstration of an optical buffer (a device that can store light) that preserves the 'quantumness' of a light beam. A quantum buffer can be used to create quantum signals on demand, synchronize quantum bits and lengthen the distance over which a quantum-communication system can work. But Marino *et al.* went beyond this mark and demonstrated that their quantum buffer can preserve a precious quantum resource called entanglement — a quantum correlation that two objects can share even if physically separated. What's more, the researchers showed that entangled quantum images (images that exhibit quantum correlations in transverse dimensions) can also be preserved in the device.

In 1935, Albert Einstein, Boris Podolsky and Nathan Rosen (EPR) wrote a paper[2] discussing what is now termed 'continuous-variable entanglement'. While trying to refute the idea that, in quantum mechanics, the wavefunction description of a physical system was complete, they ended up setting the stage for quantum information. They showed — correctly, even though they didn't agree with the results — that quantum mechanics predicts correlations between particles that can't be accounted for by classical mechanics.

Many researchers have contributed to our understanding of entanglement[3–7]. However, knowledge of continuous entanglement, as envisaged by EPR, was given a boost by Reid and Drummond[8]. Instead of looking at the correlations between the position and momentum of two particles, they considered those that can exist between 'squeezed' light beams[9]. Squeezed light beams obey the Heisenberg uncertainty relation (for example, the position and momentum of a particle can be known precisely, but not both at the same time) between light's electric-field components, which are referred to as quadratures and are analogues of the position and momentum.

The vacuum (minimum-uncertainty) state of light can be graphed in this quadrature representation as a circle at the origin of an x–y plane — one quadrature on the x-axis and the other on the y-axis. The product of the dimensions (diameter squared) of the circle is equal to the minimum uncertainty. The simplest squeezed beam is called a squeezed vacuum, and is represented by an ellipse — rather than a circle — at the origin. It is described as 'squeezed' because, to obey Heisenberg's uncertainty principle, one axis of the circle is compressed at the expense of the elongation of the other.

In EPR squeezed beams, two light beams exhibit correlations in their quadratures that cannot be explained classically. For example, under the right conditions, the sum of their phases (which describe the positions of the light's peaks and troughs) and the difference in their intensities can be less than the possible values of two uncorrelated beams. Marino *et al.*[1] used the inseparability criterion, which depends on the joint quadratures, as a reference for quantifying the degree of entanglement between two light beams. This criterion consists of calculating bounds (inseparability bounds) below which the beams, even if both squeezed, cannot be described independently[10,11].

The quantum optical buffer used in Marino and colleagues' experiment is a gas of warm rubidium atoms. After creating two entangled, squeezed light beams (using a scheme known as four-wave mixing), one of the beams was sent to the quantum buffer. An additional pump laser beam, near Raman resonance with the squeezed light beam, amplified this beam and changed the gas's index of refraction to create a region of steep dispersion (rapid change in index as a function of wavelength of light). This steep dispersion caused the squeezed light beam to travel more slowly in the optical buffer than it would have done in free space. Marino *et al.* measured delays of up to 32 nanoseconds and showed that the two beams were inseparable (entangled) — not spatially but in quantum terms — both before and after one of them passed through the buffer.

But Marino and colleagues' buffer has one primary limitation. The amplification and delay of the squeezed light beam depend on the power of the pump laser: the higher the power, the more gain and delay. But crucially,

this comes at the cost of a reduction in entanglement. With a cap on the power, the maximum delay for which entanglement can be maintained is fixed. However, this limitation should not be viewed as a 'no go' for squeezed-light buffers because there are ways around it, which several groups are pursuing.

Quantum information is not necessarily restricted to using qubits (two-state systems), because a photon is not limited to being in two states: it can be in more than two and in a superposition of many states at one time. Marino *et al.* demonstrated that their quantum buffer could also delay quantum images, which can be thought of as a superposition of many transverse states — those encoding the spatial information of the light beam. Although others have recently achieved all-optical buffering of quantum signals[12,13] and classical images[14–17], Marino *et al.* are the first to demonstrate the preservation of a quantum image (an image of the reduced Planck constant, $\hbar$), noting the transverse correlations between their two light beams. With the hope of dramatically increased information-storage capacity, it is likely that quantum images will play an important part in the future of quantum-information processing. ■

John C. Howell is in the Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627, USA.
e-mail: howell@pas.rochester.edu

1. Marino, A. M., Pooser, R. C., Boyer, V. & Lett, P. D. *Nature* **457,** 859–862 (2009).
2. Einstein, A., Podolsky, B. & Rosen, N. *Phys. Rev.* **47,** 777–780 (1935).
3. Bohm, D. *Quantum Theory* (Prentice-Hall, 1951).
4. Bell, J. S. *Physics* (Long Island City, 1965).
5. Clauser, J. F., Horne, M. A., Shimony, A. & Holt, R. A. *Phys. Rev. Lett.* **23,** 880–884 (1969).
6. Aspect, A., Grangier, P. & Roger, G. *Phys. Rev. Lett.* **47,** 460–463 (1981).
7. Ekert, A. *Phys. Rev. Lett.* **67,** 661–663 (1991).
8. Reid, M. D. & Drummond, P. D. *Phys. Rev. Lett.* **60,** 2731–2733 (1988).
9. Ou, Z. Y., Pereira, S. F., Kimble, H. J. & Peng, K. C. *Phys. Rev. Lett.* **68,** 3663–3666 (1992).
10. Duan, L.-M., Giedke, G., Cirac, J. I. & Zoller, P. *Phys. Rev. Lett.* **84,** 2722–2725 (2000).
11. Simon, R. *Phys. Rev. Lett.* **84,** 2726–2729 (2000).
12. Hètet, G. *et al. Opt. Express* **16,** 7369–7381 (2008).
13. Broadbent, C. J., Camacho, R. M., Xin, R. & Howell, J. C. *Phys. Rev. Lett.* **100,** 133602 (2008).
14. Kasapi, A., Jain, M., Yin, G. Y. & Harris, S. E. *Phys. Rev. Lett.* **74,** 2447–2450 (1995).
15. Jain, M., Merriam, A. J., Kasapi, A., Yin, G. Y. & Harris, S. E. *Phys. Rev. Lett.* **75,** 4385–4388 (1995).
16. Camacho, R. M., Broadbent, C. J., Ali-Khan, I. & Howell, J. C. *Phys. Rev. Lett.* **98,** 043902 (2007).
17. Residori, S., Bortolozzo, U. & Huignard, J. P. *Phys. Rev. Lett.* **100,** 203603 (2008).

DIAGNOSTICS

The prostate-cancer metabolome

Cory Abate-Shen and Michael M. Shen

Non-invasive detection and prognostic evaluation of cancer represents a formidable challenge. Studies of the entire metabolite composition of cells promise advances towards this objective for prostate cancer.

Prostate cancer — the most frequently diagnosed cancer in men — is a main cause of morbidity and mortality. Although it can be effectively diagnosed using a combination of digital rectal examination and measuring the levels of the enzyme PSA in the blood serum, its features are notoriously variable among patients. This makes it difficult to identify those at greatest risk of their disease progressing to advanced stages. Improved screening approaches to stratify patient populations could therefore allow selective application of more aggressive treatment strategies. On page 910 of this issue, Sreekumar *et al.*[1] report applying metabolomics to discover biomarkers that could potentially be used for non-invasive diagnosis and prognostic evaluation of prostate cancer.

So, what is metabolomics? In the post-genomic era, cancer researchers survey not only the total genetic composition of cells and tissues (genome), but also their overall collection of gene transcripts (transcriptome), proteins (proteome) and other 'omes'. The goal of such investigations is to decipher molecular signatures that distinguish tumours from normal tissues. The metabolome is the latest 'ome' to capture the imagination of researchers, and corresponds to the metabolite content of cells or tissues. Analyses of the various 'omes' are complementary, but determining the metabolite content of cancer cells — cancer metabolomics — is particularly attractive because it can provide an accurate read-out of tumours' cellular physiology and biochemical activity, with individual metabolites representing end-points of the molecular pathways perturbed by events in the other 'omes'[2,3] (Fig. 1a).

In practice, however, analysis of the metabolome is highly complex because of the broad range of detectable metabolites. Recent advances in the application of metabolomics to cancer have depended on technical improvements in the separation of various metabolites using enhanced spectrometric methods, as well as the use of systems-biology approaches for data analysis. But, despite several studies[4–6] that have made initial inroads into using applied metabolomics for the detection or prognostic evaluation of kidney and colon cancers, until now cancer metabolomics has remained in its infancy.

Sreekumar and colleagues[1] have embarked on the identification of metabolites that distinguish normal (benign) prostate from prostate cancer and its advanced metastatic form. Among the 1,126 metabolites they isolated from samples of prostate tissue, serum and urine, they identify 87 that distinguish prostate cancer from benign prostate tissue. Of these, six metabolites were of particular interest because their levels were even higher in metastatic cancer. The authors further pursued one of these metabolites as a possible biomarker for cancer progression: sarcosine — a derivative of the amino acid glycine — was the metabolite they chose owing to its elevated levels during cancer progression as well as its possible links to the disease mechanisms.

In vivo, sarcosine (*N*-methylglycine) is generated by the enzymatic transfer of a methyl group from *S*-adenosylmethionine to glycine (Fig. 1b). This reaction is catalysed by the enzyme glycine-*N*-methyltransferase (GNMT), which is expressed at high levels in the mammalian liver, exocrine pancreas and prostate. GNMT plays a central part in modulating the cellular pool of *S*-adenosylmethionine, which is the main methyl donor for several essential reactions that regulate gene expression and protein activity; these include cytosine methylation of DNA, lysine methylation of histone proteins and arginine methylation of histones and other proteins. In an earlier study[7], this group[1] found that the levels of the methyltransferase enzyme

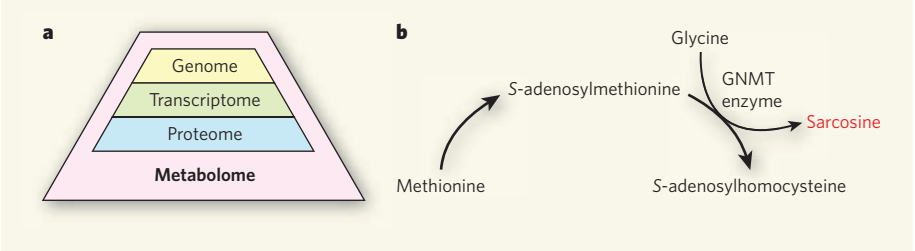


Figure 1 | Metabolomics in the bigger picture. a, Metabolomics — the relatively young approach among the analysis of various biological 'omes' — is particularly valuable for cancer research because individual metabolites represent end-points of the molecular pathways that are perturbed by other 'omes' such as the genome, transcriptome and proteome. **b**, Sreekumar *et al.*[1] used metabolomics to show that metabolites such as sarcosine, which is generated from the amino acids methionine and glycine through the enzymatic activity of GNMT, could be potential biomarkers for diagnosis and prognostic evaluation of prostate cancer, in conjunction with other established diagnostic procedures.

EZH2, which transfers methyl groups to lysine 27 of histone H3, are higher during the progression of prostate cancer and other tumours. Their latest data[1] are suggestive of a transcriptional link between cancer progression and GNMT activity, through the binding of both the androgen receptor and the oncogene *ERG* to the promoter sequence of the *GNMT* gene in tumour cells.

Unexpectedly, Sreekumar *et al.* not only show that sarcosine is a biomarker for prostate-cancer progression, but also provide evidence, using cells maintained in culture, for its functional role in regulating the features of cancer cells. They find that the addition of sarcosine to benign prostate epithelial cells promotes invasive properties in these cells, whereas lowering GNMT levels in a prostate-cancer cell line reduces its invasiveness.

By contrast, however, a previous study[8] showed that mice lacking *GNMT* develop liver cancer with age. Moreover, in a significant proportion of human prostate cancers, *GNMT* undergoes a phenomenon called loss of heterozygosity — in which one copy of the gene is lost — and the expression of this gene was documented to decrease with prostate-cancer progression[9]. Reconciling these earlier findings with those of Sreekumar *et al.* is necessary to determine the overall significance of sarcosine levels in assessing cancer progression.

At present, the greatest value of metabolomic approaches seems to be for the development of non-invasive screening procedures that can be used for effective cancer diagnosis and prognosis. Notably, Sreekumar and colleagues[1] show that sarcosine levels in urine have a modest but significant predictive value for prostate-cancer diagnosis; this suggests that assessment of metabolite levels in urine might be an appropriate screening tool when applied together with examination of PSA levels and other approaches for monitoring disease progression. Furthermore, the authors identify several other metabolites that are more readily detected in cancer and metastases than sarcosine, although with no obvious mechanistic link to disease progression; these metabolites might therefore be more suitable for predictive screening tests.

It is not known whether metabolome changes similar to those Sreekumar *et al.* observe in prostate cancer occur in other tumours. It will also be of interest to learn how environmental factors such as diet (including intake of methionine — the precursor of *S*-adenosylmethionine) may affect the metabolome profile and thus the usefulness of metabolomic analysis in cancer screening. As a starting point, however, Sreekumar and colleagues' observations suggest that metabolomics has a promising future in aiding cancer diagnosis and treatment. ■

Cory Abate-Shen is in the Departments of Urology and of Pathology, and Michael M. Shen is in the Departments of Medicine and of Genetics and Development, Herbert Irving Comprehensive Cancer Center, Columbia University College of Physicians and Surgeons, New York, New York 10032, USA.
e-mails: cabateshen@columbia.edu; mshen@columbia.edu

1. Sreekumar, A. *et al. Nature* **457,** 910–914 (2009).
2. Griffin, J. L. & Shockcor, J. P. *Nature Rev. Cancer* **4,** 551–561 (2004).
3. Spratlin, J. L., Serkova, N. J. & Eckhardt, S. G. *Clin. Cancer Res.* **15,** 431–440 (2009).
4. Denkert, C. *et al. Mol. Cancer* **7,** 72 (2008).
5. Ippolito, J. E. *et al. Proc. Natl Acad. Sci. USA* **102,** 9901–9906 (2005).
6. Kind, T. *et al. Anal. Biochem.* **363,** 185–195 (2007).
7. Varambally, S. *et al. Nature* **419,** 624–629 (2002).
8. Martínez-Chantar, M. L. *et al. Hepatology* **47,** 1191–1199 (2008).
9. Huang, Y.-C. *et al. Clin. Cancer Res.* **13,** 1412–1420 (2007).

SOLID-STATE CHEMISTRY

Boron charged under pressure

John S. Tse

Crystal-structure prediction methods and diffraction data show that a newly discovered form of boron is partially ionic. This is the first time such a structure has been observed for any elemental solid.

The concept of atomic structure — ordered arrangements of atoms, such as the arrays found in crystalline solids — is central to an understanding of the properties of matter. The most successful technique for determining crystal structures is X-ray diffraction. But this method doesn't always work, especially for unusually complex structures, or if the resolution of the technique is limited by the extreme conditions required for the crystals to exist. In these cases, an independent method for predicting plausible structures is desirable to assist the interpretation of experimental diffraction patterns. Such a method is extremely difficult to realize in practice. But reporting in this issue (page 863), Oganov *et al.*[1] demonstrate the use of first-principles calculations, in combination with diffraction data, to determine the crystal structure of a new, stable form of boron that forms at high pressure.

The theoretical prediction of crystal structures has long been a challenging problem. Recently, there has been a resurgence of interest in applying first-principles calculations to predict structures, without having any prior information about the arrangement of atoms. These include *ex nihilo* approaches[2] (which generate structures at random, and then use computational methods to 'relax' the structures into low-energy arrangements) and genetic algorithms[3] (which start with a population of random structures, from which low-energy structures 'evolve' during subsequent rounds of changes that simulate natural selection). An exciting aspect of Oganov and colleagues' work[1] is that it used the latter approach to successfully predict a thermodynamically stable structure.

The authors' method[3] for predicting crystal structures builds on previous pioneering work[4] that predicted the structures of clusters of atoms. The process starts with a population of randomly generated parent structures, from which offspring structures evolve in heredity and mutation operations. In heredity operations, two low-enthalpy structures are selected and sliced up, and slices from each structure are then combined to produce offspring (Fig. 1). Mutation operations create distortions to the unit cell (the fundamental repeat unit of a crystalline structure) and/or to the atomic positions of a parent structure. After several rounds of operations, the lowest-energy crystal structure should emerge.

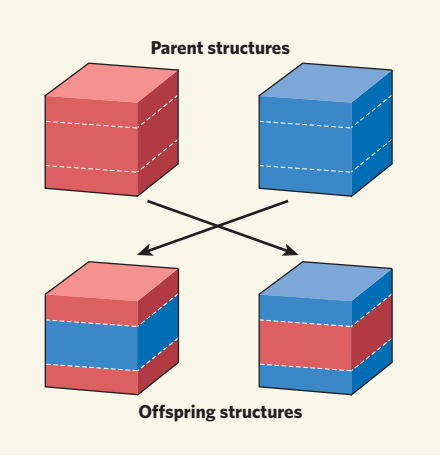


Figure 1 | Heredity operations in crystal-structure prediction. Genetic algorithms are used to predict crystal structures of compounds. These involve heredity operations, whereby models of the unit cells of approximate, likely structures (parent structures) are cut into slices, and slices from different parents are combined to produce offspring structures. After several heredity and mutation operations (which generate distortions to unit cells), low-energy structures of the compound under investigation should emerge. Oganov *et al.*[1] use genetic algorithms to determine the crystal structure of a newly identified form of boron.

Boron provides a tough test for the authors' predictive method because of its chemical and structural complexity. Discovered in 1808, the original samples of boron were later found to contain only 60–80% of the element, and

it wasn't until 1909 that a 99% pure sample was isolated. So far, at least 16 polymorphs (structural forms) of boron are known, yet the ground-state structure of boron is still controversial. Oganov *et al.*[1] report a new form of boron that adds to this *mélange* — a polymorph that is stable at high pressure (18–89 gigapascals). So what is its structure?

Working with a 'quenched' sample of the material (one that had been recovered to ambient pressure without disrupting its structure), the authors first obtained some basic information about the unit cell of the polymorph. They then applied their prediction technique to a range of possible atomic arrangements, and successfully obtained a thermodynamically stable structure that reproduced the observed diffraction pattern of the material. The structure is made up of icosahedral clusters of 12 boron atoms (B_{12}) and of pairs of boron atoms (B_2), arranged in a cubic lattice (see Fig. 1b on page 864). It is both gratifying and exciting that the correct structure was predicted, despite its complexity.

A surprising finding is that the bonding between the B_{12} clusters and the B_2 pairs is partially ionic, as suggested by the infrared spectrum of the polymorph and supported by the theoretical calculations. It was generally believed that the effect of pressure on a covalently bonded compound was to weaken the bonds, and so enhance electron mobility. For example, it was suggested[5] that a simple compound, such as solid hydrogen (an insulator consisting of H_2 molecules, in which electrons are localized within covalent bonds), would transform under extreme pressure to a metal (containing mobile, delocalized electrons). More recent experimental and theoretical results[6], however, suggest that spontaneous polarization of H_2 bonds might occur under pressure, so that partial charges develop on the atoms of each molecule — electrons situated in the interstitial region would leave positive charge on the atoms.

Oganov *et al.*[1] report the first observation of an elemental solid that has some ionic structure — the B_{12} clusters have a partial negative charge, whereas the B_2 units have a partial positive charge. The build-up of negative charge on the B_{12} cluster is stabilized by sharing the negative charge (electron) among the molecular orbitals of the B_{12} units. This effect, combined with electrostatic interactions between the B_2 and the B_{12} units, helps to stabilize the ionic complex. This picture, if correct, is different from the proposed spontaneous electronic polarization in dense hydrogen, and may be unique to boron.

Oganov and colleagues' results prove the efficiency and reliability of their prediction technique, but there is still some room for improvement. Simply mixing and matching planar slices of parent structures in the heredity operation isn't necessarily the best procedure for finding low-energy structures — slices that have periodic, wave-like geometries might be more efficient[7]. The mutation operation could (and should) also be improved.

Furthermore, there have been considerable successes in predictions of crystal structures that are based solely on randomly generated structures[8] — the advantage of this approach being that the sampling of possible structures is completely unbiased. Such lack of bias might prevent the evolution process from predicting structures that lie in a local energy minimum of the potential energy surface that describes all the possible arrangements of atoms (rather than finding structures of the lowest possible energy).

Oganov and colleagues' paper[1] describes a crucial first step towards a truly *ab initio* method for predicting crystal structures. Perhaps just as importantly, it broadens our knowledge of the possibilities for chemical bonding in elemental solids, and adds yet another chapter to the bizarre history of boron. ■

John S. Tse is in the Department of Physics and Engineering Physics, University of Saskatchewan, 116 Science Place, Saskatoon, Saskatchewan S7N 0K4, Canada.
e-mail: john.tse@usask.ca

1. Oganov, A. R. *et al. Nature* **457,** 863–867 (2009).
2. Pickard, C. J. & Needs, R. J. *Phys. Rev. Lett.* **97,** 045504 (2006).
3. Oganov, A. R. & Glass, C. W. *J. Chem. Phys.* **124,** 244704 (2006).
4. Deaven, D. M. & Ho, K. M. *Phys. Rev. Lett.* **75,** 288–291 (1995).
5. Wigner, E. & Huntington, H. B. *J. Chem. Phys.* **3,** 764–770 (1935).
6. Edwards, B. & Ashcroft, N. W. *Nature* **388,** 652–655 (1997).
7. Abraham, N. L. & Probert, M. I. J. *Phys. Rev. B* **73,** 224104 (2006).
8. Pickard, C. J. & Needs, R. J. *Nature Phys.* **3,** 473–476 (2007).

DEVELOPMENTAL BIOLOGY

Birth of the blood cell

Momoko Yoshimoto and Mervin C. Yoder

Could it be that mouse fetal liver cells and adult bone-marrow blood cells originate from a subset of cells that line the blood vessels in the embryo? Several lines of evidence suggest that this is indeed the case.

During embryonic development, haematopoietic stem cells, which give rise to blood cells, and endothelial cells, which line blood vessels, both form from the mesodermal germ-cell layer; but exactly how is debatable. On the one hand, a controversial, century-old theory proposes that both haematopoietic and endothelial cells arise from a mesoderm-derived common precursor called a haemangioblast. On the other hand, a younger theory proposes that haematopoietic stem cells (HSCs) form from a subset of early endothelial cells known as haemogenic endothelium. The relationship between haemangioblasts and haemogenic endothelium has never been resolved. In this issue, however, three papers[1–3] clarify the potential relatedness and significance of these cell types.

The concept of the haemangioblast initially arose from observations that, in the chick yolk sac, a mesoderm-derived cell can give rise to both primitive red blood cells and endothelial cells. Moreover, the finding that, in the mouse yolk sac, the formation of blood islands — aggregates of blood cells and endothelium — requires the expression of specific genes such as *Flk-1* provided further support for the relatedness of the two cell lineages. But the strongest evidence for the existence of haemangioblasts was obtained after the development of the BL-CFC *in vitro* assay. This assay allows clonal (single-cell) analysis of blast colony-forming cells (BL-CFCs), which are derived from differentiating mouse embryonic stem (ES) cells[4]. By definition, BL-CFCs can directly form both haematopoietic and endothelial cells, and are therefore the closest detectable equivalent of the theoretical haemangioblasts .

When ES-cell-derived Flk-1-expressing (Flk-1$^+$) mouse cells are grown in culture, characteristic BL-CFC colonies appear that consist of an aggregate of non-adherent blood cells overlying an adherent layer of endothelium. This observation, together with insights[4,5] into the molecular regulation of the development and differentiation of colonies that emerge from a BL-CFC (blast colonies), has been enlightening. Nonetheless, little has become clear about the cellular events that herald the generation of blood cells from BL-CFCs.

Lancrin *et al.*[1] (page 892) used time-lapse photography to analyse the sequence of cellular events required for the formation of mature blast colonies from cultured Flk-1$^+$ cells. They find that these colonies form in two stages. First, after 36–48 hours of 'plating' Flk-1$^+$ cells for growth in culture, the cells form tightly adherent clusters. Subsequently, round, non-adherent cells appear, which then proliferate to complete the formation of mature blast colonies. Among the adherent cell clusters at 48 hours, a transient cell population expressing various endothelial (but not mesodermal or BL-CFC) markers appear, displaying the potential to form haematopoietic cells. From this population, both primitive blood-cell colonies eventually form, characterized by their expression of the embryonic version of the haemoglobin protein, together with definitive blood-cell colonies expressing adult haemoglobin.

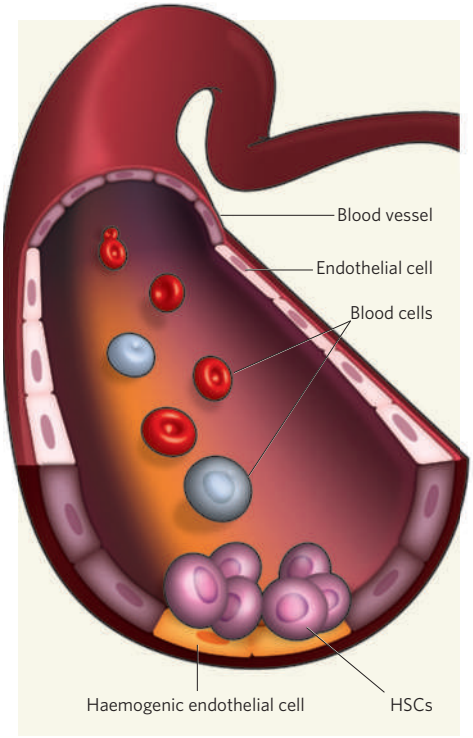


Figure 1 | Relationship between endothelial cells and blood cells. Endothelial cells line the inside of blood vessels. During mouse embryonic development, a subset of these cells, known as haemogenic endothelial cells, seems to give rise to haematopoietic stem cells (HSCs) and their progenitors, such as those that seed the fetal liver and the adult bone marrow[1–3].

Lancrin and colleagues' observations suggest that haematopoietic progenitor cells arise from haemangioblasts through a haemogenic endothelial intermediate — the first linear pathway to resolve, at least *in vitro*, the relationship between haemangioblasts and haemogenic endothelium (Fig. 1). But do these findings alter the definition of the haemangioblast? To answer this question, we must learn more about the fate of haemogenic endothelial cells following the birth of blood cells. For example, it will be interesting to assess whether haemogenic endothelial cells that give rise to primitive blood cells differ from those that produce definitive blood cells.

In many species, HSCs appear as clusters attached to the endothelium that lines the ventral wall of the abdominal aorta during embryonic development; this observation has long implicated the endothelium as the source of developing blood cells. Indeed, when endothelial cells isolated from mouse embryos are grown in culture, a subset has the potential to develop into mature blood cells such as erythroid, myeloid and/or lymphoid cells[6]. But this and other evidence is only indirect[7], and direct proof of haematopoietic progenitor cells emerging from individual endothelial cells has been lacking.

Eilken *et al.*[2] (page 896) tracked the fates of all cells (more than 6,500) generated from individually plated mouse ES-cell-derived mesoderm cells using time-lapse microscopy. Their detailed analysis of the resulting colonies indicates that 1.2% of the colonies display properties of adherent endothelial cells, and that one or more endothelial cells in a colony directly give rise to non-adherent haematopoietic cells. The authors also directly isolated primary endothelial cells with haemogenic potential from early mouse embryos. They therefore demonstrate that haemogenic endothelial cells not only can be generated *in vitro* from ES cells but are also naturally present in mouse embryos. But the question that these authors[2] and Lancrin *et al.*[1] did not address is whether HSCs emerge directly from haemogenic endothelial cells *in vivo* during mouse development.

In the developing mouse embryo, the transcription factor Runx1 is required for the formation of HSCs and their progenitors. In fact, Runx1 has been considered necessary for the emergence of HSC clusters from the haemogenic endothelium[8]. Chen *et al.*[3] (page 887) show that, within the endothelium, Runx1 expression is indeed essential for the formation of HSCs and their progenitors over a period of roughly 3 days during mouse embryonic development (embryonic days 8.25–11.5). Furthermore, in agreement with another recent report[9], they show that most fetal liver cells and adult bone-marrow cells originate from the haemogenic endothelium.

Together, these studies[1–3] provide substantial evidence that HSCs and their progenitor cells that populate the fetal liver and adult bone marrow originate from the differentiated endothelium that resides in the functional vasculature of the mouse conceptus. The focus can now turn to determining the intriguing molecular mechanisms involved, which might differ between the various embryonic sites of blood-cell production[10]. What's more, translation of this knowledge to the human system could be of great assistance in generating human HSCs from human ES cells, either by direct cell reprogramming[11] or indirectly through induced pluripotent stem cells. ■

Momoko Yoshimoto and Mervin C. Yoder are in the Herman B. Wells Center for Pediatrics Research, Indiana University, Indianapolis, Indiana 46202, USA.
e-mail: myoder@iupui.edu

1. Lancrin, C. *et al. Nature* **457,** 892–895 (2009).
2. Eilken, H. M., Nishikawa, S.-I. & Schroeder, T. *Nature* **457,** 896–900 (2009).
3. Chen, M. J., Yokomizo, T., Zeigler, B. M., Dzierzak, E. & Speck, N. A. *Nature* **457,** 887–891 (2009).
4. Choi, K., Kennedy, M., Kazarov, A., Papadimitriou, J. C. & Keller, G. *Development* **125,** 725–732 (1998).
5. Huber, T. L., Kouskoff, V., Fehling, H. J., Palis, J. & Keller, G. *Nature* **432,** 625–630 (2004).
6. Nishikawa, S. I. *et al. Immunity* **8,** 761–769 (1998).
7. de Bruijn, M. F. *et al. Immunity* **16,** 673–683 (2002).
8. Yokomizo, T. *et al. Genes Cells* **6,** 13–23 (2001).
9. Zovein, A. C. *et al. Cell Stem Cell* **3,** 625–636 (2008).
10. Dzierzak, E. & Speck, N. A. *Nature Immunol.* **9,** 129–136 (2008).
11. Gurdon, J. B. & Melton, D. A. *Science* **322,** 1811–1815 (2008).

EVOLUTION

Unnatural selection

Nils Chr. Stenseth and Erin S. Dunlop

Fishing and hunting by humans are the main causes of mortality in many populations of wild animals. The consequence is that large and rapid changes occur in certain characteristics that far exceed changes due to other agents.

Harvesting by humans exerts tremendous pressure on wild populations, resulting in both ecological and evolutionary change. Writing in *Proceedings of the National Academy of Sciences*, Darimont *et al.*[1] demonstrate the consequences of harvesting on changes that occur in observable traits (phenotypes) in populations experiencing one of three different types of selection.

Using a meta-analysis of previously published data, the authors compared the rates of phenotypic change in 40 populations subject to human harvesting with the rates seen in 20 systems that experienced selection from natural forces (for example, Darwin's finches) and with the rates in 25 systems that experienced other human disturbance (for example, pollution). The human-harvested organisms included fish, ungulates, invertebrates and even plants, examples of these being respectively sockeye salmon, bighorn sheep, marine snails and gingseng.

The recorded rates of change in harvested populations were astonishingly high, outpacing changes arising from natural agents by 300% and those from other anthropogenic causes by 50%. The changes were both rapid and dramatic in magnitude: declines in morphological traits (such as body size) averaged about 20%, and shifts in life-history traits (such as reproductive age) averaged nearly 25%. Darimont *et al.* conclude with a thoughtful discussion about the effects of such changes on harvestable biomass, and whether populations will show a similarly swift return to previous trait distributions if recovery is a goal of resource management.

Darimont and colleagues[1] argue that large

NATURE|Vol 457|12 February 2009

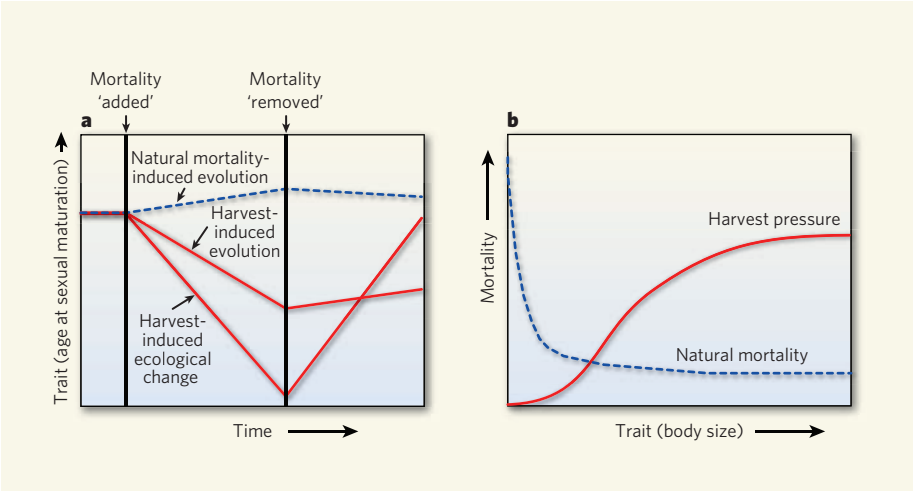


Figure 1 | Genetic (evolutionary) and plastic (ecological) responses to mortality from human harvesting and natural forces. **a**, Harvesting-induced evolution (of, for example, age at sexual maturation) is often faster than evolution caused by natural mortality. Ecological changes induced by harvesting pressure occur more quickly than evolutionary changes, but the latter are expected to be longer lasting when harvesting-induced mortality is removed. Natural mortality could also cause a rapid ecological response, but this is not shown. **b**, Harvesting pressure — from such activities as trawling or trophy hunting — operates on a trait such as body size in the opposite direction to mortality from natural sources[7].

and rapid phenotypic changes have profound implications, regardless of whether they are due to genetic (evolutionary) or plastic (purely ecological) responses. Although a plastic response will alter some aspects of population dynamics, genetic changes have further effects that, from a management perspective, suggest that it is essential to identify them. Most notably, a population is expected to recover more slowly from genetic than from plastic changes (Fig. 1a), and genetic change might even be irreversible[2]. Given the high harvesting pressure on larger and older individuals, the result might be 'juvenescence' of populations, possibly leading to increased variability in abundance and reduced genetic variability[3]. The evolution of traits related to growth and reproduction could also lead to reductions in the yield and profits derived from exploited populations.

Why do harvesting-induced changes outpace those caused by other sources? The response, R, of a trait to selection is determined by the breeder's equation, $R = h^2S$, where h^2 is heritability of the trait (the ratio of additive genetic variance to phenotypic variance) and S is the selection differential (a measure of the strength of selection)[4]. Assuming that heritabilities are not systematically different among the various populations included in Darimont and colleagues' survey, it seems that the selection differential generated from harvesting is higher than that from other sources. An obvious reason is that the mortality rates produced from human fishing and hunting are often far greater than natural mortality rates.

Most notably, mortality arising from commercial fishing can exceed natural mortality by as much as 400% (ref. 5). In the years immediately preceding the collapse of stocks of Atlantic cod off Newfoundland and Labrador, the likelihood[6] of an individual of harvestable size being caught was as high as 89%. The strength of the selective pressures generated by such high rates of mortality no doubt contributes to the observed faster rates of change generated by human harvesting than by other sources of trait change.

Besides resulting in high rates of mortality, the selective pressures induced by harvesting typically act in the opposite direction to those caused by natural sources of mortality[7] (Fig. 1b). Natural mortality rates of young or small individuals are usually much higher than for old or large individuals, for example because predators are often limited by the size of their gape. Human predators are different, usually preferring to hunt and capture large prey for sport, profit or sustenance. Mortality from trawl fishing increases with a fish's body size because the smallest fish can escape through the trawl mesh. Similarly, many recreational fisheries have size limits that ban the taking of small individuals, and trophy hunters often preferentially target large game with, for example, impressive antlers.

Natural mortality is, of course, always acting in the background. As the different sources of selection often oppose one another[7], this might produce intermediate rates of change. But harvesting mortality is usually far stronger than natural mortality. So selection from the former would be expected to swamp that due to the latter, and rates of change may still be fairly high and faster than those induced by natural mortality alone.

One commonly observed trend, and one for which Darimont *et al.*[1] obtained data, is that fishing causes the evolution of sexual maturation at younger ages and smaller sizes[8]. In the absence of fishing, it pays to be big, both because you can escape predation and because there is a positive relationship between body size and fecundity. Small females that reproduce early in life will still gain some reproductive advantage, but it will not be as strong as for those females that reproduce when large and highly fecund. In the presence of fishing, being large becomes a liability, as large individuals are targeted by the fishery. In this case, females that delay maturation until they are larger could have a total lifetime reproductive fitness of zero, making the effect of selection much stronger, because the fish will be killed before having the chance to reproduce. Therefore, natural selection favouring large size and delayed maturation is often weaker than the selection induced by fishing, possibly contributing to the disparate rates of change reported by Darimont and colleagues[1].

One implication of this study[1] is that natural-resource management should include consideration of evolutionary responses to fishing and hunting. Evolutionary-impact assessments can be used for this purpose, where predictive models are used to evaluate the potential evolutionary impact of alternative management actions[8]. The underlying model should include both ecological and evolutionary parameters (and the interaction between them), as well as 'utility' measures such as yield and profit. The evidence for harvesting-induced evolution remains under debate, and the relative importance of ecological and evolutionary contributions to observed change remains contentious[9–12]. Nonetheless, a precautionary approach[13] requires that management strategies be designed under the assumption that harvesting-induced evolution might occur. The potential costs of ignoring that possibility are severe — most notably, the resulting changes may be difficult or impossible to reverse. ■

Nils Chr. Stenseth is at the Centre for Ecological and Evolutionary Synthesis, Department of Biology, University of Oslo, PO Box 1066, Blindern, N-0316 Oslo, Norway. Erin S. Dunlop is at the Ontario Ministry of Natural Resources, 2140 East Bank Drive, Peterborough, Ontario K9J 7B8, Canada.
e-mail: n.c.stenseth@bio.uio.no

1. Darimont, C. T. *et al. Proc. Natl Acad. Sci. USA* **106,** 952–954 (2009).
2. Stenseth, N. C. & Rouyer, T. *Nature* **452,** 825–826 (2008).
3. Anderson, C. N. K. *et al. Nature* **452,** 835–839 (2008).
4. Falconer, D. S. & Mackay, T. F. C. *Introduction to Quantitative Genetics* (Longman, 1996).
5. Mertz, G. & Myers, R. A. *Can. J. Fish Aquat. Sci.* **55,** 478–484 (1998).
6. Myers, R. A. *et al. Mar. Ecol. Prog. Ser.* **138,** 293–308 (1996).
7. Carlson, S. M. *et al. Ecol. Lett.* **10,** 512–521 (2007).
8. Jørgensen, C. *et al. Science* **378,** 1247–1248 (2007).
9. Hilborn, R. *Fisheries* **31,** 554–555 (2006).
10. Law, R. *Mar. Ecol. Prog. Ser.* **335,** 271–277 (2007).
11. Browman, H. I. *et al. Science* **320,** 47 (2008).
12. Kuparinen, A. & Merilä, J. *Science* **320,** 47–48 (2008).
13. Food and Agriculture Organization *Precautionary Approach to Capture Fisheries and Species Introductions* FAO Tech. Guidelines for Responsible Fisheries No. 2 (FAO, 1996).

For more on Darwin, see www.nature.com/darwin.

ANALYTICAL CHEMISTRY

Reactions assayed by magnets

David E. Bergbreiter

Does it float or sink? And to what extent? The answers to these questions can be used to follow the course of chemical reactions on solid supports, and are obtained simply by using two magnets, a salt solution and a ruler.

A generation of children is familiar with the concept of levitation from the adventures of Harry Potter, and could be forgiven for thinking it unremarkable. Scientists will know of real-world examples — reports of levitating water-droplets[1], cells[2] and even frogs[3] — but most of them probably think of it as nothing more than a curiosity. It's unlikely that anyone imagined that levitation could be used as a routine analytical tool. But in the *Journal of the American Chemical Society*, Mirica *et al.*[4] show that magnetic levitation is a simple, viable way to follow the course of chemical reactions that occur on solid, diamagnetic supports, such as those routinely used by chemists to prepare peptides, nucleic acids and other organic molecules. (Diamagnetic materials are those that create a weak magnetic field in opposition to an external magnetic field.)

The authors use their levitation procedure to analyse beads made from insoluble polymers known as Merrifield resins — named after their inventor Robert Merrifield, who won a Nobel prize in 1984 for his discovery[5]. These polymers were originally used as ion-exchange resins[6] (which trap ions in analytical and purification procedures), but they were widely adopted by synthetic chemists as solid supports for reactions after Merrifield described their use for making peptides. Beads of Merrifield resins quickly became standard tools for preparing biologically important molecules; indeed, they enabled the biotechnology revolution. Nowadays, they are used in all manner of organic syntheses, not only as supports for reactions (for example, in combinatorial chemistry), but also as supports for reagents and catalysts[7] (because such resin-bound compounds are easily removed by filtration from reaction mixtures in solution, greatly simplifying purification procedures).

Although polymer-supported reactions have been widely used for almost 50 years, the procedures available to monitor them are surprisingly limited[8]. Gel-phase nuclear magnetic resonance spectroscopy is the most useful technique, but this requires sophisticated instruments and is not always practical as a real-time assay. Colorimetric assays — which, as the name suggests, qualitatively monitor the amount of colour formed in a test reaction — can sometimes provide a rapid, qualitative assessment of the extent of a resin-bound reaction. But up to now, no analytical procedure could provide an almost instantaneous, qualitative and general assay of the progress of solid-supported reactions. Mirica and colleagues' technique[4] promises to change that.

The authors' method requires only a test tube, a pair of commercially available magnets and an inexpensive solution of gadolinium chloride ($GdCl_3$) in water (Fig. 1). When polymeric beads are suspended in the solution in a vertical magnetic field, the height at which they float depends on the balance between the gravitational and magnetic forces acting on them. This in turn depends on the density of the polymer, which changes when reactions occur on the beads. Mirica *et al.* thus monitored changes in the 'floating height' of beads to follow the progress of chemical reactions on the beads.

The authors' method for tracking the course of solid-supported reactions is considerably simpler than previously available methods, which require more-complex instruments and are more limited in scope. Furthermore, the levitation analysis is sensitive enough to detect even slight changes in density between beads. Mirica *et al.* thus showed that their method could be used not only to determine when a reaction is complete, but also to follow reaction kinetics. Differences in floating heights could also be used to separate mixtures of beads that have different densities[9].

Although the levitation procedure described is clearly useful, its advantages and limitations still have to be defined. For example, the authors show that, at the mid-point of a reaction, a broader range of floating heights is obtained than at the beginning or end of the reaction. This could complicate some analyses. It also seems that, for the best results, the beads should all be the same size (which isn't always the case), and that traces of solvent adsorbed to the beads can affect the outcome of the assay. On the other hand, opportunities exist to optimize the technique, by designing synthetic processes or resins that maximize changes in density caused by reactions, or by devising an automated version of the method that uses optical sensors to monitor changes in the heights of floating beads — currently, the authors measure the heights by hand using a millimetre-scale ruler.

Chemistry evolves through the invention of new reactions or catalysts, and less commonly through the discovery of new ways of doing things. Mirica and colleagues' gravimetric levitation analysis falls into the second category. Scientists who have the seemingly simple — but practically difficult — problem of monitoring the course of reactions on solid supports now have a new, easy way to do it. ■

David E. Bergbreiter is in the Department of Chemistry, Texas A&M University, PO Box 30012, College Station, Texas 77843-3012, USA.
e-mail: bergbreiter@tamu.edu

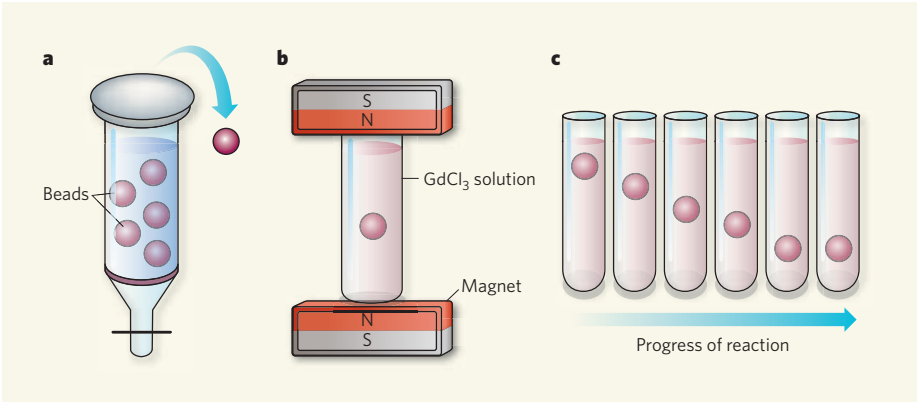
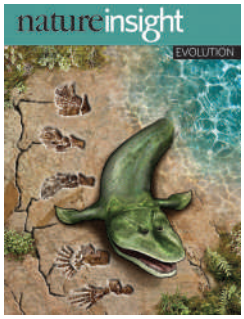


Figure 1 | Sink or levitate. a, Many synthetic reactions are performed using polymeric beads. Each bead is typically 75–150 micrometres in diameter, so many hundreds of beads are used in each reaction. (Beads are not shown to scale). **b**, When added to a solution of gadolinium chloride ($GdCl_3$) in a magnetic field, such beads levitate. The height at which they levitate depends on their density, which varies during the course of a reaction. **c**, Mirica *et al.*[4] use this effect to follow the progress of solid-supported reactions. In their method, beads are removed at intervals from a reaction, and changes in the height at which they levitate are observed. When no further change in levitation height is observed for successive beads, the reaction is complete.

1. Ikezoe, Y., Hirota, N., Nakagawa, J. & Kitazawa, K. *Nature* **393,** 749–750 (1998).
2. Coleman, C. B. *et al. Biotechnol. Bioeng.* **98,** 854–863 (2007).
3. Simon, M. D. & Geim, A. K. *J. Appl. Phys.* **87,** 6200–6204 (2000).
4. Mirica, K. A., Phillips, S. T., Shevkoplyas, S. S. & Whitesides, G. M. *J. Am. Chem. Soc.* **130,** 17678–17680 (2008).
5. http://nobelprize.org/nobel_prizes/chemistry/laureates/1984/index.html
6. Alexandratos, S. D. *Ind. Eng. Chem. Res.* **48,** 388–398 (2009).
7. Toy, P. H. & Shi, M. *Tetrahedron* **61,** 12025 (2005).
8. Sabatino, G., Chelli, M., Brandi, A. & Papini, A. M. *Curr. Org. Chem.* **8,** 291–301 (2004).
9. Winkleman, A. *et al. Anal. Chem.* **79,** 6542–6550 (2007).

nature insight

EVOLUTION



Cover illustration
Fossilized limbs based on illustrations by K. Monoyios (Univ. Chicago, Illinois). Artwork by N. Spencer

Darwin200

EVOLUTION

The articles in this Insight testify to the success of Charles Darwin's theory of descent with modification by means of natural selection, carefully detailed in his book *On the Origin of Species* almost 150 years ago. The most striking aspect of the theory is its simplicity. Given heritable variation, a superabundance of offspring, and environmental change, natural selection must happen, and evolution will follow. The natural world can be explained without invoking pre-existing germs, essential life forces, the great chain of being, Ptolemaic epicycles or a prime mover.

This simplicity has meant that the theory has always accommodated new discoveries — the general theme of this Insight. In Darwin's day, nothing was known about genetics or the mathematical basis of natural selection. But such discoveries have only made the theory stronger.

Simplicity also makes for longevity. The theory of natural selection has had its ups and downs, but today we are not celebrating, for example, the 280th anniversary of the birth of the great experimental scientist Lazzaro Spallanzani. Why not? He saw in his results confirmation of the theory of preformation: that the essence of organisms is stamped in the egg, and all that is needed is for the pre-existing germ to unfold. This theory died with Spallanzani, overcome by better observation — and by a theory with no preconditions. No one subscribes to the theory of preformation now, whereas natural selection continues to evolve.

To be sure, Darwin's first ideas now seem dated, but he winnowed them over decades, stripping them of any archaic clutter to reveal a modern clarity of purpose on which biologists have continued to build.

So, as we toast the bicentenary of Darwin's birth today, we can be sure that Darwin's name will be familiar to our descendants (however modified) for centuries to come, whereas those of Spallanzani and many others — so great in their day — will succumb to the inevitable flip side of evolution: extinction.

Henry Gee, Senior Editor,
Rory Howlett, Consultant Editor

OVERVIEW

REVIEWS

nature insight

NATURE|Vol 457|12 February 2009|doi:10.1038/nature07889

Natural selection 150 years on

Mark Pagel[1,2]

The theory of evolution by natural selection has prospered in its first 150 years and provides a consistent account of species as highly adapted and rare survivors in the struggle for existence. It now faces the challenge of finding order in the evolution of complex systems, including human society.

Happy 200th birthday to Charles Darwin, whose theory of evolution by natural selection will be 150 later this year. Darwin had been quietly amassing support for his ideas for nearly 20 years when in 1858 he discovered that Alfred Russel Wallace was thinking along similar lines. In the following year, Darwin's 50th, he finally published the theory in *On the Origin of Species*[1].

It was a theory marked out for controversy from birth. Darwin's claim was that natural selection could explain the diversity of all life on Earth, from bacteria to barnacles, orchids, fish, lizards, mushrooms, pine trees, elephants, bananas, stick insects and bonobos. Provocatively, given the religious views that prevailed at the time, it was a materialistic account of nature. There were no guiding hands, no grand plans, no 'great chain of being'. According to Darwin[1], the varieties of forms that are observed arose because "As many more individuals of each species are born than can possibly survive; and as, consequently, there is a frequently recurring struggle for existence, it follows that any being, if it vary however slightly in any manner profitable to itself, under the complex and sometimes varying conditions of life, will have a better chance of surviving, and thus be naturally selected. From the strong principle of inheritance, any selected variety will tend to propagate its new and modified form."

It is an idea of remarkable simplicity. Complex organisms emerge from the gradual accumulation of successive modifications, each one of which improves the bearer's chance of surviving in the struggle for existence. But Darwin would go further, yielding up what must be one of the most daring hostages to fortune in the history of science, declaring: "if it could be demonstrated that any complex organ existed which could not possibly have been formed by numerous, successive, slight modifications, my theory would absolutely break down."

Darwin has been spared the embarrassment he courted. Natural selection is an idea that has survived and propagated in the competitive environment of the mind. There are no serious alternative explanations for the existence and diversity of species. Most would agree that the theory has been immensely successful. It has been capable of explaining — as adaptations that improve survival and reproduction — many features of organisms: their shapes, sizes and colours; why they live so long or have such short lives; their reproductive habits, diet and mating choices; the so-called organs of perfection, such as eyes or brains; and even social behaviours, explaining parental care, altruism, sibling rivalry and aggression. It has been influential in understanding viral, bacterial and other microbial evolution, and in explaining the evolutionary changes that take place at the level of DNA itself.

The theory's application has not been limited to biology. In the past few decades, it has jumped to many other disciplines, with papers on natural selection appearing in the scientific journals of, among others, anthropologists, sociologists, philosophers, economists, mathematicians and statisticians, demographers, physicists and even surgeons, yielding a wide and 'long tailed' distribution (Fig. 1).

Given the continued success of the theory of evolution by natural selection, it is useful to step back and take stock of what it indicates about the nature of the biotic world. Does all life descend, as Darwin thought, from a long unbroken line of ancestors, or have the various forms of life evolved independently? How creative a force is natural selection? Are all human traits adaptations? Are humans just one of many, possibly arbitrary, outcomes of the evolutionary process, or would the same kinds of species and adaptations arise if the tape of evolution were re-run? Is there an unlimited variety of species?

In this Overview, I review where the theory stands on these core questions, which cut across the hundreds of specific topics of evolutionary investigation. This list of questions is by no means comprehensive, and many researchers will find their favourite question ignored. But the answers that emerge here and in the articles that follow in this Insight provide a glimpse into a process — natural selection — that is at once conservative and severe, yet innovative and exacting, yielding organisms that are often surprisingly well adapted from among a relatively small number that manage to wriggle or claw their way through its sieve.

Humans are among the survivors, but the answers pondered here do far more than tell us about ourselves. They set a benchmark for the kinds of phenomenon that the many other disciplines that make use of Darwinian ideas might expect to see in their observations. At the end of this Overview, I briefly discuss how natural selection is approaching two growing areas of research into what are becoming known as 'complex adaptive systems'[2]: one relates to the rules controlling the organization of bodies such as those of humans; the other seeks to understand human societal evolution.

Descent with modification

Darwin used the word 'evolve' just once in *On the Origin of Species*, and even then he waited until the final word of the book. Instead, Darwin wrote of 'descent with modification'. For many at the time, this meant species descending in the 'great chain of being', with God at its peak. This was a progressive view of life, in which each level was seen as more evolved than the one below. Humans occupied the earthly pinnacle of the chain, having supplanted the simpler monkeys.

But the only illustration that Darwin put in *On the Origin of Species* shows evolution following a branching, or tree-like, process, in which the different branches of the tree represent the separate forms of life (Fig. 2). Successful routes through the tree reach to the present, whereas those that fail result in extinction. This tree banished in a single stroke the idea of progress in evolution; instead, all extant life is equally evolved. A 'lowly' slime mould has been evolving as long as you have, even if you regard yourself as more complex. Slime moulds are good at being slime moulds, and they are better at being slime moulds than you are or than a giraffe is. Giraffes, in turn, are good at being giraffes and better at being them than fish or bananas are, and so on. What natural selection makes species good at is 'propagating', something that is now recognized as transmitting copies of their genes into future generations.

[1]School of Biological Sciences, University of Reading, Reading RG6 6AJ, UK. [2]The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA.

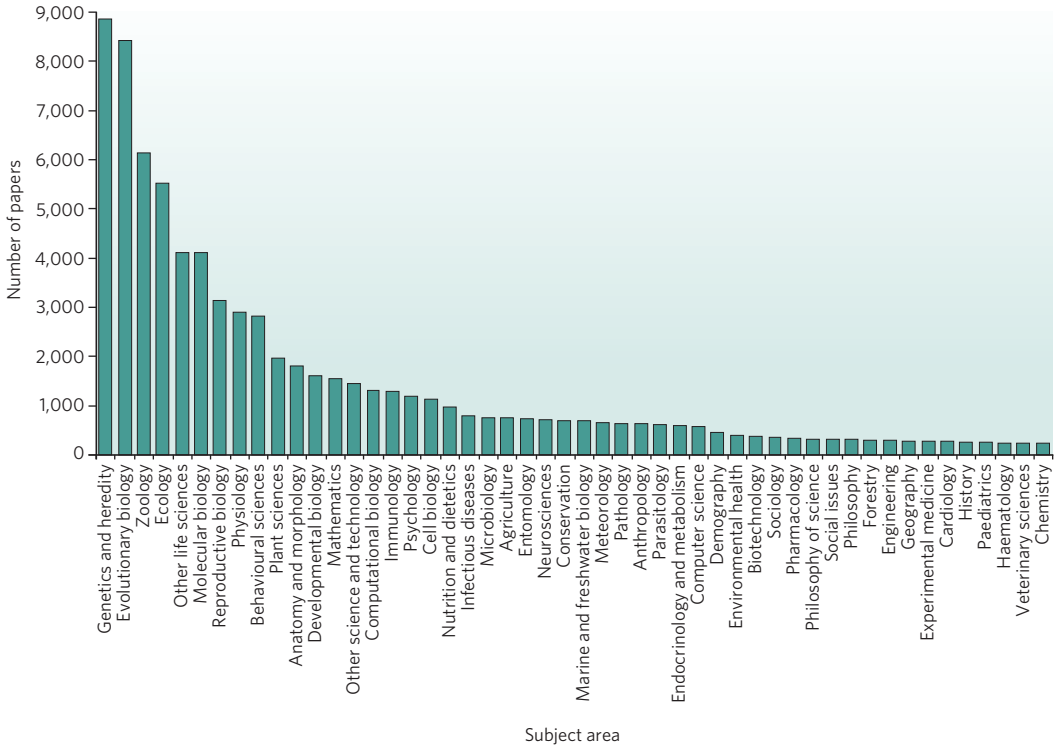


Figure 1 | The long tail of natural selection in scholarly work. The number of papers that include the term 'natural selection' in their title, abstract or keywords, recorded separately for subject areas as identified by the ISI Web of Knowledge. Data are derived from a search on 'natural selection' in November 2008, yielding 14,232 hits over all years. This figure underestimates the number of papers that investigate topics in evolution, being limited only to those that include the search term. The figure reports the top 50 subject areas from over 140. Subject areas are not mutually exclusive, so the total of the quantities on the *y* axis exceeds 14,232.

Darwin concluded that humans were in the ape part of the tree, and this led to some ribaldry among his supporters and opponents. At a debate at the Oxford Union in 1860, the bishop of Oxford, Samuel Wilberforce, supposedly asked Thomas Henry Huxley, a staunch supporter of Darwin, whether it was through his grandfather or his grandmother that he claimed descent from a monkey[3]. Huxley's reply can be read at leisure[3]. What matters is that genetics has shown that Darwin was right. The nearest living genetic relative to humans is the chimpanzee, with the common ancestor of these species existing some 6–8 million years ago. Both species belong to a group known as the hominids, comprising humans and their extinct relatives, along with the great apes (chimpanzees, gorillas and orang-utans). Its evolution has been bushy or tree-like, not a progressive line leading inevitably to us[4,5]. More than 20 species of hominin — the subgroup that includes humans and the immediate fossil relatives of humans — have been identified from fossils, including australopithecines such as Lucy (*Australopithecus afarensis*) and *Homo* species such as Neanderthals. During part of history, up to five different hominin species have coexisted[5]. Now just one survives: *Homo sapiens*.

Darwin's branching idea would prove correct for more species than just the apes. Evolutionary biologists refer to groups of species that derive from a common ancestor not shared with any other species as 'monophyletic'. Were species to march backwards in time[6] down the branches of Fig. 2, all of the various mammals would meet up at a common ancestor to the exclusion of non-mammals; the birds would meet up at their own unique common ancestor; so too would all of the vertebrates, a huge group that includes all animals with a backbone (such as birds and mammals). On page 812 of this Insight, Graham Budd and Maximilian Telford report on the monophyly of a different group: the arthropods. These comprise the majority of all animal species on Earth, including insects and crustaceans, and they all derive from a single common ancestor.

Budd and Telford place this group on the tree of life next to the worms rather than the vertebrates (as though worms and arthropods are represented by b^{14} and f^{14} at the top left of Fig. 2 and vertebrates are represented by a^{14}, q^{14} and p^{14}), forming a group known as Ecdysozoa[7]. For years, it was thought that arthropods grouped with vertebrates. Budd and Telford's placement has implications for which species are the most relevant model organisms for understanding humans: should researchers be studying worms or fruitflies? The existence of Ecdysozoa suggests that both are equally relevant (or irrelevant). Moving deeper in the tree of life, past the animals, Darwin would be further vindicated by the fact that all animals share an ancestor with the plants and fungi. And then, going back further, all species eventually share an ancestor some 3.5 billion years ago, somewhere among the prokaryotes (the bacteria and Archaea), from which all life on Earth descends. No one knows if life evolved more than once, but all extant life seems to have a single source.

Tinkerer or perfectionist

The size and details of monophyletic groups illustrate an important feature of life. Rather than designing each species from scratch, as an engineer might, evolution is conservative, using the same designs over and over. Darwin recognized, as the comparative anatomist Geoffroy Saint-Hilaire had before, that the hands of moles, horses, porpoises and bats all used the same bones. Such traits, derived from common ancestral traits, are called homologies. Neil Shubin, Cliff Tabin and Sean Carroll (see page 818) trace the deep lines of conservation of many traits in ancient homologies hundreds of millions of years old. Eyes, for example, have evolved independently perhaps dozens of times, but the gene encoding the protein PAX6 is implicated in one way or another in nearly every known instance of a light-sensitive organ appearing in an animal's body. This includes the light-sensitive tissues of some shellfish, as well as human eyes.

Pervasive monophyly and homology raises the question of whether humans actually are adapted. Is natural selection a mere tinkerer or a perfectionist? Arthur Cain was unequivocal in a 1964 article memorably entitled 'The Perfection of Animals'[8]. Although it was not one of Cain's examples, the sonar of insectivorous bats (acquired 70 million years ago) has features that match those of the best sonar produced by human engineers[9]. Others were less sanguine about adaptation. In 1979, Stephen Jay Gould and Richard Lewontin attacked what they saw as an 'adaptationist programme' in evolutionary biology that uncritically treated all traits as adaptations[10].

Which view is correct? Not everything is an adaptation: human blood just happens to be red, and human chins might be relics of the way the human jaw develops. But the weight of evidence suggests that it is probably wise not to bet against natural selection. The struggle for existence means that traits have to pay their way. The traits observed now probably improve an animal's chances of surviving and propagating, and those traits that do not will tend to be lost. For example, fish that have adapted to life in dark underwater caves lose the ability to see.

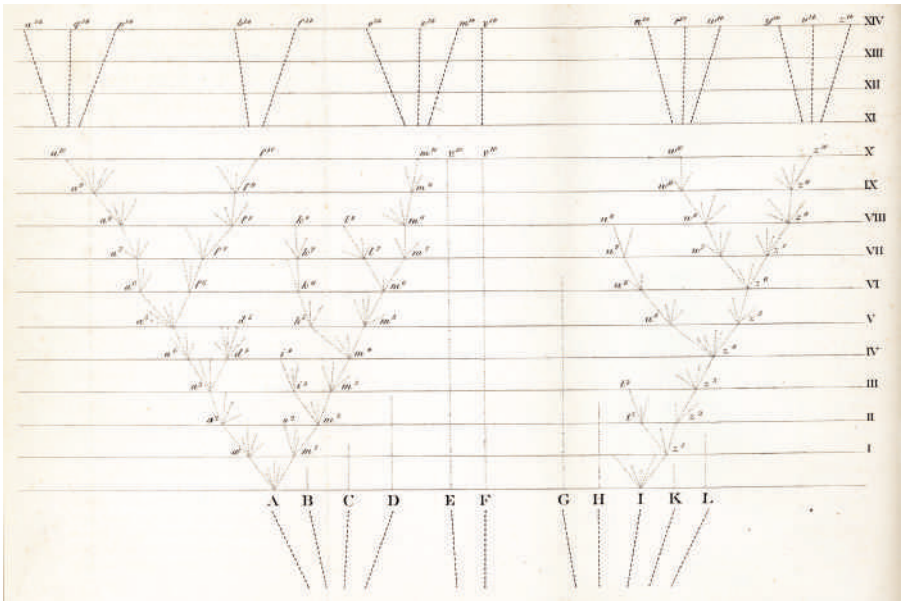


Figure 2 | Darwin's tree. A reproduction of the only figure Darwin included in *On the Origin of Species*[1]. The diagram shows Darwin's conception of evolution as a branching, or tree-like, process. Lines that reach to the top of the diagram represent extant, or surviving, species. Lines that end farther down are lineages that have gone extinct, revealing Darwin's hunch that extinction rates are high. The two large groups, or clades, of branching species issuing from common ancestors A and I are monophyletic: that is, each of these groups derives from a common ancestor that is not shared with any other species. Common ancestors within these clades are indicated by letters with superscript numerals where lineages converge. Lineages associated with capital letters reach further into the past. (Figure reproduced from ref. 1.)

One difficulty in studying evolution is that such traits tend to evolve over vast periods of time. But there are ways of seeing evolution in action. On page 824, Angus Buckling, R. Craig Maclean, Michael Brockhurst and Nick Colegrave review the field of 'experimental evolution', in which microorganisms are observed in test tubes over tens of thousands of generations. Rates of evolution are high enough in these studies to test in real time and under controlled conditions many of the basic ideas of adaptation and natural selection.

Such studies show regular and reliable adaptation as microbial lines diversify. They emphasize the importance of chance mutations in producing differences among individuals (Darwin's 'slight variations'). They also demonstrate that the ability to adapt can itself evolve; for example, organisms can adaptively change the rate at which they produce new random variants. More adaptive evolution is observed at the genetic level than is expected given the redundancies in the DNA code[11,12]. Most evolutionary changes are small, but they can sometimes sweep through populations rapidly, producing sudden bursts of evolution[13–15]. Crucially for Darwin's theory, later forms can be shown to survive and reproduce better than earlier forms when compared in the same environment.

Variation

A paradox of Darwinism is that variety is maintained among individuals of the same species. Repeated bouts of natural selection might be expected to use up genetic variation by selecting for the best features, making evolution grind to a halt. Many factors can contribute to variation being maintained, but experimental evolution studies (see page 824) and the analysis of innovations (see page 818) point to two fundamental and often overlooked effects.

First, as organisms have increased in complexity over time, natural selection has had more to work with. For example, it often co-opts pre-existing genetic regulatory circuits, fashioning them to take on new roles (see page 818). Second, as time passes, diversity seems to beget more diversity by creating different ways for species to make a living. For example, the first species to appear is unlikely to be a predator, but the second might be. This idea matches the human cultural experience of domesticating dogs. In just a few thousand years, humans created varieties ranging from chihuahuas to Great Danes. Each new breed seems to have suggested another, and dogs did not run out of the genetic variation that humans fancied.

Contingency

How repeatable is evolution? If the tape of evolution were replayed, would there once again be turkeys and tomatoes, rattlesnakes, the plague and humans? If the species observed now are the chance outcomes of many contingent historical events, there is hardly anything special about any particular outcome of evolution, including humans. Gould championed such 'contingency' in evolution, pointing out that humans might not be here were it not for the fortuitous extinction of the dinosaurs, which preceded humans on this planet[16] but died out as a result of a meteor impact. Perhaps so, but recent work[17] suggests that mammals were alive, well and vigorously diversifying before the dinosaurs became extinct.

Studies of evolution on islands can provide direct evidence relating to the question of contingency. In effect, the tape of evolution gets played out independently on each island. And if the islands are similar enough, they can be taken as replicates of each other. On page 830, Jonathan Losos and Robert Ricklefs discuss adaptive radiations on islands, and use as an example species of *Anolis* lizard on relatively isolated islands in the Caribbean. They report that evolutionary diversification is often similar on the different islands, yielding the same set of habitat specialists, each adapted to use a different part of the vegetation. Buckling *et al.* (see page 824) report similar repeatability in independent lines of microorganisms in experimental evolution studies. Contingency does not seem to be the pervasive force that Gould suspected. In other instances, islands present unique opportunities for natural selection — and unique faunas, such as those of New Zealand and Madagascar, are the result.

Speciation

Even the Pope now accepts the reality of evolution among individuals of a single species, but he and many others with religious beliefs draw the line at speciation. It is a clever ploy, tipping one's hat to science by giving evolution a minor role while reserving new species for the providence of creation. Speciation can require thousands of years, making it difficult to observe. It does not help that Darwin said little about the origin of species in *On the Origin of Species*. In addition, closely related species are expected to be similar to each other, and yet some yawning gaps exist. For example, the rock hyrax, a small creature, is thought to be the closest living relative of the elephants.

Regarding these gaps, on page 837, David Reznick and Ricklefs explain that if the rates of extinction of species are high (Fig. 2), then gaps would be expected among the surviving species. No one knows precisely what historical extinction rates have been, but a reasonable guess is that they could be nearly as high as speciation rates; extant species are a rarefied club of survivors. Many of these survivors might be responsible for the gaps by outcompeting the similarly shaped and sized species to which they were closely related. The hominins might be a good example: as brain size and intelligence increased in this group, the less imaginative species could have been driven extinct. The close human cousins the Neanderthals might be the most recent case.

But can speciation be observed? Yes. Field biologists have witnessed speciation in action among several plant, amphibian, bird and fish species[18]. What matters for Darwin's theory is that the process by which populations of interbreeding individuals split into two non-interbreeding populations follows straightforward routes of natural selection. One good example is that females of a cichlid fish species vary genetically in their preferences for males of red and blue colours. Biologists are witnessing red males occupying the lower depths of Lake Victoria, in Africa, and females with matching preferences are following them. These mating preferences are causing this single species to split into two[19].

Ways of life

Here is a fundamental logical challenge to the idea of adaptation. If there were an infinite number of potential niches, or ways of life, for organisms to fill, it would be difficult to argue that species are adapted, because anything could do well. Darwin drew inspiration for his theory of natural selection by observing the effects of 'artificial selection' by farmers and others during the domestication of their crops and animals. But there is nothing artificial about artificial selection. On page 843, Michael Purugganan and Dorian Fuller show that domestication is just another way of life, in this case a kind of plant–animal mutualism, and their conclusions are relevant to the question of infinite niches.

Some ants can herd aphids as though they are cattle or can domesticate fungi, and termites also practice domestication[20]. But humans have taken it to new heights. There is an abundance of human-selected wheat, oilseed, barley, maize and rice genes on the planet, not to mention cows, sheep and chickens. When humans do the domesticating, it is clear what the 'selector' is hoping to achieve. Purugganan and Fuller suggest that humans have frequently sought to improve just two characteristics — plant germination and ease of harvesting — and they find that a relatively small number of groups of genes lie behind most instances of successful plant domestication.

Their conclusions invite the speculation that nature — like its human counterpart — admits far fewer combinations of genes, and by inference kinds of species, than are possible. There are, for example, no banana-eating snakes or flying monkeys, no species between humans and chimpanzees, nor between rock hyraxes and elephants. If the varieties of life are severely constrained, the competition to occupy them will be fierce. Natural selection emerges as a severe and vigilant master.

A theory moving with the times

Some complex systems can acquire information about the environment and respond to it, becoming 'complex adaptive systems'[2]. Genomes exhibit this behaviour on an evolutionary timescale, whereas the bodies of most organisms and human societies do so in real time. How does this kind of adaptive complexity evolve, and is it correct to think of it as serving the interests of a single entity?

As evolutionary studies enter the genomic era, biologists are finding out more and more about how genes combine with a surprisingly large range of genetic regulatory mechanisms to produce the complex systems that are the phenotypes of organisms such as ourselves[21,22]. The genes of most multicellular organisms can often be counted on to 'pull together' because they have the same route of reproduction and live or die together in the same body, although there is scope for conflict[23]. How natural selection favours phenotypes that are reliable and robust to outside influences, but still able to adapt to new circumstances, will be central to understanding development.

The question of whose interests are served is sharpened once natural selection is allowed to venture into the realms of cultural and societal evolution. The big complex adaptive system that is human society is leaky: there are many different independent replicators — both biological individuals and cultural elements — each potentially with its own strategies for survival and reproduction. Should human society be viewed as a vehicle for the combined, cumulative effects of these replicators, rather than as a replicating system in its own right? If so, what rules govern which vehicles are successful, and do they bear any relationship to those for biological phenotypes?

There is a growing sense, for example, that human languages have adapted to human minds[24]. Humans have domesticated languages: languages show features related to how they are used and to society[15,25,26], and this probably enhances their survival. Language might also to some degree have domesticated humans[24]. It might have a regulatory role in human society not unlike that of gene regulation[27], and this may have enhanced human survival. Much the same could be said about the interactions between humans and the varieties of religion, art and music, topics that interested Darwin[28]. The ability of natural selection to keep up with the times as more and more questions are asked shows that, far from being old at 150, Darwin's theory still has a spring in its step. ■

1. Darwin, C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (John Murray, 1859).
2. Gell-Mann, M. *The Quark and the Jaguar: Adventures in the Simple and Complex* (Freeman, 1994).
3. Lucas, J. R. Wilberforce and Huxley: a legendary encounter. *Hist. J.* **22,** 313–330 (1979).
4. Tattersall, I. Paleoanthropology: the last half-century. *Evol. Anthropol.* **9,** 2–16 (2000).
5. Stringer, C. & Andrews, P. *The Complete World of Human Evolution* (Thames & Hudson, 2005).
6. Dawkins, R. *The Ancestor's Tale: A Pilgrimage to the Dawn of Life* (Houghton Mifflin, 2004).
7. Aguinaldo, A. M. A. *et al.* Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387,** 489–493 (1997).
8. Cain, A. J. The perfection of animals. *Biol. J. Linn. Soc.* **36,** 3–29 (1964).
9. Dawkins, R. *The Blind Watchmaker* (Longman, 1986).
10. Gould, S. J. & Lewontin, R. C. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. R. Soc. Lond. B* **205,** 581–598 (1979).
11. Fay, J. C., Wyckoff, G. J. & Wu, C. I. Testing the neutral theory of molecular evolution with genomic data from *Drosophila*. *Nature* **415,** 1024–1026 (2002).
12. Smith, N. G. C. & Eyre-Walker, A. Adaptive protein evolution in *Drosophila*. *Nature* **415,** 1022–1024 (2002).
13. Lenski, R. E. & Travisano, M. Dynamics of adaptation and diversification — a 10,000-generation experiment with bacterial populations. *Proc. Natl Acad. Sci. USA* **91,** 6808–6814 (1994).
14. Pagel, M., Venditti, C. & Meade, A. Large punctuational contribution of speciation to evolutionary divergence at the molecular level. *Science* **314,** 119–121 (2006).
15. Atkinson, Q., Meade, A., Venditti, C., Greenhill, S. & Pagel, M. Languages evolve in punctuational bursts. *Science* **319,** 588 (2008).
16. Gould, S. J. *Wonderful Life* (Norton, 1989).
17. Bininda-Emonds, O. R. P. *et al.* The delayed rise of present day mammals. *Nature* **446,** 507–512 (2007).
18. Coyne, J. A. Speciation in action. *Science* **272,** 700–701 (1996).
19. Seehausen, O. *et al.* Speciation through sensory drive in cichlid fish. *Nature* **455,** 620–626 (2008).
20. Schulz, T. R. & Brady S. G. Major evolutionary transitions in ant agriculture. *Proc. Natl Acad. Sci. USA* **105,** 5435–5440 (2008).
21. Kirschner, M. W. & Gerhart, J. C. *The Plausibility of Life: Resolving Darwin's Dilemma* (Yale Univ. Press, 2005).
22. Grimson, A. *et al.* Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature* **455,** 1193–1197 (2008).
23. Burt, A. & Trivers, R. *Genes in Conflict: The Biology of Selfish Genetic Elements* (Belknap, 2006).
24. Christiansen, M. H. & Chater, N. Language as shaped by the brain. *Behav. Brain Sci.* **31,** 489–509 (2008).
25. Pagel, M., Atkinson, Q. & Meade, A. Frequency of word-use predicts rate of lexical evolution throughout Indo-European history. *Nature* **449,** 717–720 (2007).
26. Lieberman, E., Michel, J.-B., Jackson, J., Tang, T. & Nowak, M. Quantifying the evolutionary dynamics of language. *Nature* **449,** 713–716 (2007).
27. Pagel, M. Rise of the digital machine. *Nature* **452,** 699 (2008).
28. Darwin, C. *The Descent of Man, and Selection in Relation to Sex* (John Murray, 1871).

Author Information Reprints and permissions information is available at www.nature.com/reprints. The author declares no competing financial interests. Correspondence should be addressed to the author (m.pagel@reading.ac.uk).

NATURE|Vol 457|12 February 2009|doi:10.1038/nature07890

The origin and evolution of arthropods

Graham E. Budd[1] & Maximilian J. Telford[2]

The past two decades have witnessed profound changes in our understanding of the evolution of arthropods. Many of these insights derive from the adoption of molecular methods by systematists and developmental biologists, prompting a radical reordering of the relationships among extant arthropod classes and their closest non-arthropod relatives, and shedding light on the developmental basis for the origins of key characteristics. A complementary source of data is the discovery of fossils from several spectacular Cambrian faunas. These fossils form well-characterized groupings, making the broad pattern of Cambrian arthropod systematics increasingly consensual.

The arthropods are one of the most familiar and ubiquitous of all animal groups. They have far more species than any other phylum, yet the living species are merely the surviving branches of a much greater diversity of extinct forms. One group of crustacean arthropods, the barnacles, was studied extensively by Charles Darwin. But the origins and the evolution of arthropods in general, embedded in what is now known as the Cambrian explosion, were a source of considerable concern to him, and he devoted a substantial and anxious section of *On the Origin of Species*[1] to discussing this subject: "For instance, I cannot doubt that all the Silurian trilobites have descended from some one crustacean, which must have lived long before the Silurian age, and which probably differed greatly from any known animal." His interest, if not his uncertainty, was echoed repeatedly over the following 150 years, with debate over what were the closest relatives of the arthropods and over the relationships between the main constituent groups, and even doubts about whether the phylum is monophyletic[2] (that is, whether it evolved from a single common ancestor that is not shared with any other phylum).

Since the publication of *On the Origin of Species*, most data on the pattern of arthropod evolution have been obtained by studying embryos, adult morphology, and fossils, but the introduction of molecular biological data to phylogenetics and comparative developmental biology in the past 20 years has led to great insights. Gene sequences provide vast numbers of markers of phylogenetic relationships and, over the past 20 years, have redrawn many aspects of the metazoan tree of life. The comparative molecular genetic analysis of development has similarly changed the view of the evolution of developmental mechanisms and the origins of novel morphology, revealing surprising conservation and providing a complement to phylogenetic proximity for determining homology. Even the study of morphology has been changed by molecular techniques, and the palaeontological evidence has been transformed by the steady description of exceptionally well preserved fossils from the Cambrian and, increasingly, from other periods too.

In this Review, we discuss recent advances in understanding arthropod origins and relationships from the fields of molecular systematics, palaeontology, morphology and 'evo-devo'. We show that the source of Darwin's discomfort about arthropod origins, although not entirely removed, has been substantially alleviated. A new consensus is emerging about the timing of arthropod origins, as well as the relationships among arthropods (including between fossils and living taxa) and between arthropods and non-arthropods.

Arthropods are monophyletic

Arthropods encompass a great diversity of animal taxa known from the Cambrian to the present day. The four living groups — myriapods, chelicerates, insects and crustaceans — are known collectively as the Euarthropoda. They are united by a set of distinctive features, most notably the clear segmentation of their bodies, a sclerotized cuticle and jointed appendages. Even so, their great diversity has led to considerable debate over whether they had single (monophyletic) or multiple (polyphyletic) origins from a soft-bodied, legless ancestor. The application of molecular systematics to arthropods[3] in 1992, however, decisively resolved the issue in favour of monophyly[4]. In other words, many of the morphological features shared by arthropods are likely to have a single origin and to have diversified across the group.

It has long been recognized that two other living groups, the soft-bodied onychophorans (velvet worms) and the microscopic tardigrades (water bears), are close relatives of the euarthropods. All of these groups are segmented and have appendages, and they are often collectively referred to as the Panarthropoda. All of the available molecular and morphological evidence supports the idea of onychophorans and euarthropods falling into a monophyletic group or clade, but the position of the tardigrades is less clear. Although they are traditionally regarded as the closest living relatives of the euarthropods, some molecular phylogenies place them basal within the panarthropods, or even as a sister group to the nematodes[5,6], but this may be an artefact resulting from their derived and rapidly evolving genome[4,5].

Arthropods are ecdysozoans

The similarity of the arthropods to another segmented phylum, the annelid worms, has long been noted. Arthropods and annelids share several features, such as segmentation and the structure of their nervous and blood vascular systems. Since the time of Darwin, it has been widely assumed that the arthropods evolved from an annelidan ancestor. There have been notes of dissent, however[6,7], and this minority view was vindicated by the publication in 1997 of a molecular analysis of ribosomal RNA genes that introduced the concept of the Ecdysozoa[8], a clade consisting of panarthropods and a group of lesser-known worms named the Cycloneuralia, comprising the priapulids, kinorhynchs, loriciferans, nematodes and nematomorphs.

Some morphologists[9] have resisted the dissolution of the Articulata (arthropods plus annelids). Many molecular analyses using the large data sets from whole genome sequences of *Drosophila melanogaster*,

[1]Department of Earth Sciences, Uppsala University, Villavägen 16, Uppsala SE-752 36, Sweden. [2]Department of Genetics, Evolution and Environment, University College London, Gower Street, London WC1E 6BT, UK.

humans and *Caenorhabditis elegans* have grouped together vertebrates and arthropods, creating a monophyletic clade of coelomates[10–14], rather than linking arthropods and nematodes. Although annelids are not represented in these analyses, evidence for a monophyletic Coelomata is consistent with the Articulata and not with Ecdysozoa. Such extensive genomic data sets can be expected to avoid the sampling errors associated with smaller studies. The consensus amongst systematists, however, is that the rapidly evolving nematode genome causes a systematic error, known as long-branch attraction, in tree reconstruction in these studies, which artefactually places the nematode distant from the arthropods towards the root of the tree. Including data from more slowly evolving nematodes or from the priapulid worms results in unequivocal support for Ecdysozoa[5,15–18]. Furthermore, the observation of a complex characteristic of the mitochondrial genome common to all protostomes, although not addressing the Ecdysozoa versus Articulata debate, argues strongly against the Coelomata clade[4,19]. The morphological support for Ecdysozoa remains slender, although several synapomorphies have been demonstrated or can be inferred[20], including ecdysis, a trilaminate cuticle and a terminal mouth (seen in the cycloneuralians and inferred from the fossil record for the arthropods[21,22]).

Within Ecdysozoa, the relationships between the panarthropods and the cycloneuralian worms remain poorly resolved. The Cycloneuralia have been thought to be monophyletic on the basis of their shared brain anatomy[23] and the most recent phylogenomic results[5], but the possibility remains that the Cycloneuralia are paraphyletic and gave rise directly to the arthropods[24,25]. Such a reconstruction would imply that the morphological features shared by arthropods and annelids have evolved convergently or, if homologous, have been lost more than once in the Cycloneuralia. The resolution of ecdysozoan relationships will have important consequences for the reconstruction of the last common protostome ancestor.

Unexpected groupings of euarthropods

The classical view of euarthropod relationships placed hexapods (insects, diplurans, proturans and collembolans) together with myriapods in the Atelocerata (or Tracheata). The Atelocerata, in turn, were considered to be the sister group of the Crustacea, with the three classes — hexapods, myriapods and crustaceans — forming a group united by possession of a mandible and named the Mandibulata. The chelicerates were held to be a sister group to the Mandibulata. This broad grouping was challenged by Sidnie Manton[26] and fellow polyphyleticists on functional grounds, and also by many palaeontologists, who placed the Atelocerata as a sister group to the 'CCT' clade (crustaceans, chelicerates and trilobites)[27]. These alternatives have now been rejected, and the current consensus phylogenetic tree is shown in Fig. 1. Molecular systematic studies over the past ten years have convincingly removed the hexapods from their traditional position as sister group to the myriapods in the Atelocerata. The hexapods are now regarded as either a sister group to the crustaceans, the two together being regarded as the Tetraconata[28–31], or, more likely, as an in-group of the crustaceans and hence a terrestrial branch of Pancrustacea[31]. Studies of neurogenesis[32] and eye development[33] provide independent support for the Pancrustacea. Manton's old group of the Uniramia (hexapods, myriapods and onychophorans) has thus been

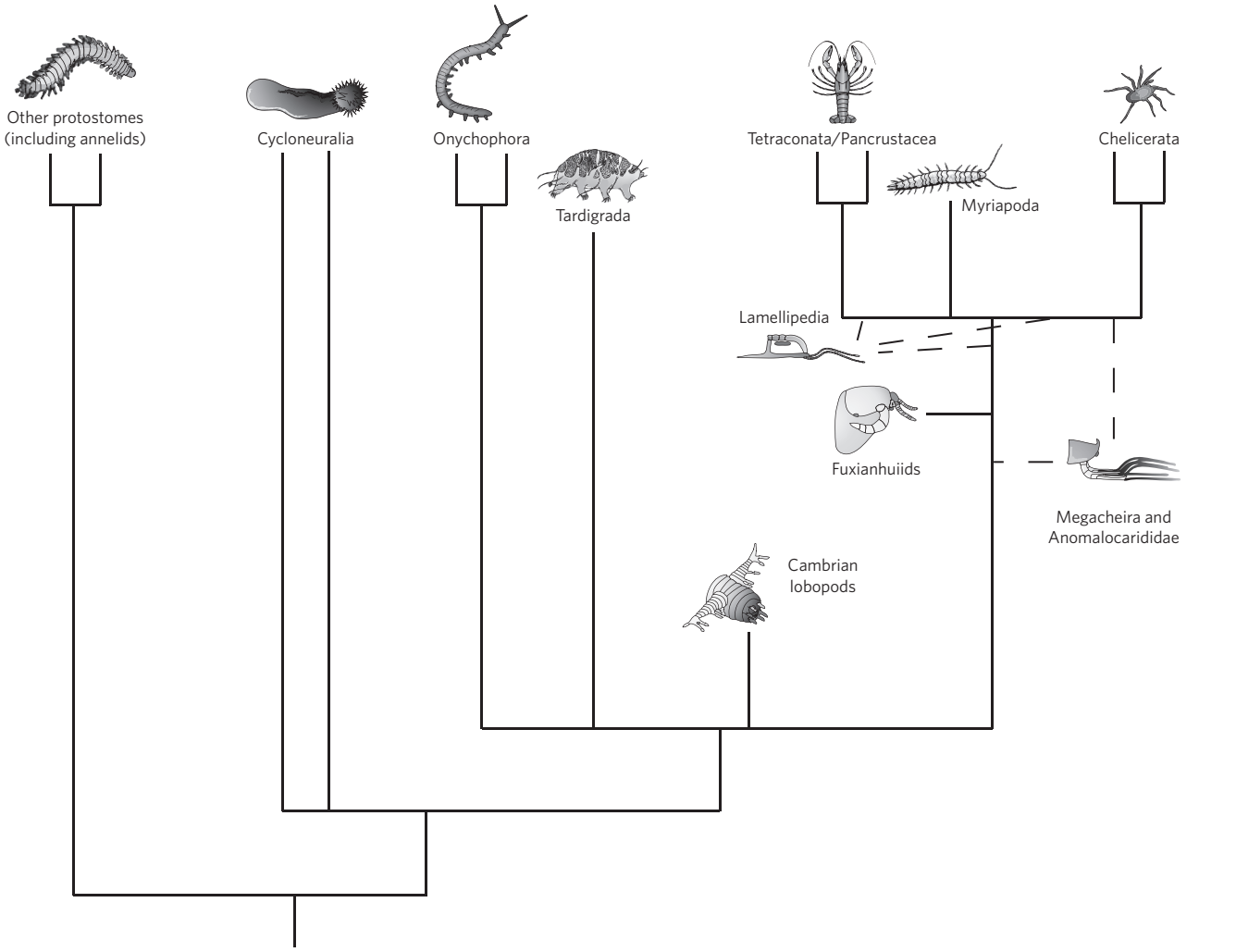


Figure 1 | Progress and problems in arthropod phylogeny. A consensus reconstruction of arthropod relationships, based on molecular, morphological and fossil data. The sister-group relationship to cycloneuralians is shown, as are the various positions suggested for upper-stem-group euarthropods. This tree differs markedly from any phylogeny that Darwin would have recognized.

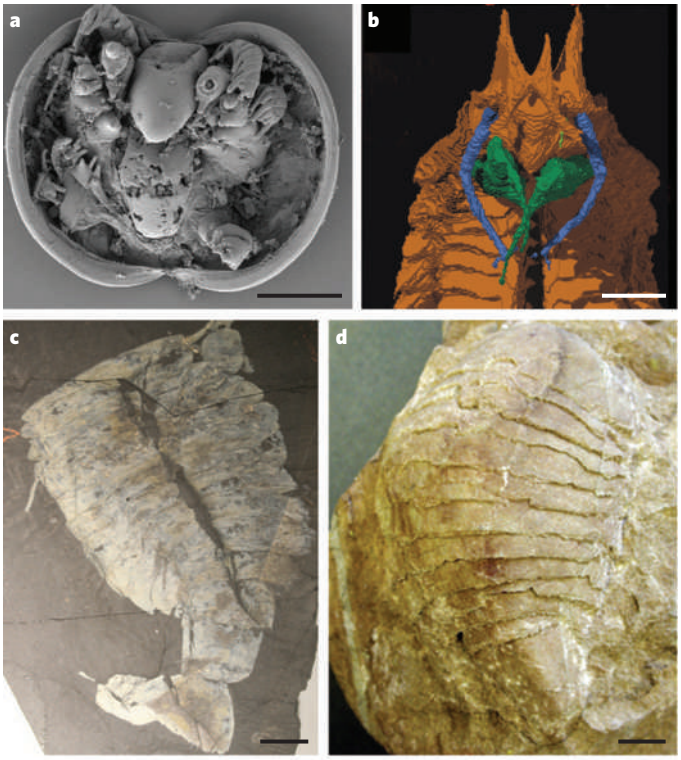


Figure 2 | Putative stem-group mandibulates and chelicerates. **a**, *Hesslandona* sp., a member of the upper stem group of the crustaceans, from the Upper Cambrian of Backeborg, Sweden. Scale bar, 50 µm. **b**, Computer reconstruction of the head region of the putative stem-group crustacean *Tanazios dokeron* from the Silurian of Herefordshire, UK. The first antenna is shown in blue, and the mandible in green. Scale bar, 1 mm. **c**, *Sidneyia inexpectans*, a possible stem-group chelicerate from the Burgess Shale (Middle Cambrian). Scale bar, 1 cm. **d**, *Paleomerus hamiltoni*, a putative stem-group chelicerate from the Lower Cambrian of Kinnekulle hill, Sweden. Scale bar, 1 cm. (Panels **a** and **c** courtesy of A. Daley (Uppsala University, Sweden); panel **b** reproduced, with permission, from ref. 40.)

thoroughly dismantled. Hexapod and crustacean relationships nevertheless remain controversial, partly because relatively little is known about potentially basal groups such as the remipedes[34] and cephalocarids[35], from either molecular or developmental perspectives. Even the monophyly of the hexapods has been questioned by mitochondrial genome analyses, and there remains the remote possibility that the Collembola represent an independent lineage of crustaceans[36,37].

Crustacean phylogeny is unique among crown-group arthropods in having a reasonable number of stem-group fossil forms, largely from the Upper Cambrian 'Orsten' fauna[38,39] (Fig. 2a) but possibly also from the Silurian[40]. Although early insect and hexapod fossils are rare and have not generally helped much in terms of reconstructing their respective stem groups, there has been a recent burst of fossil descriptions[41,42], including that of the intriguing *Tanazios dokeron* (Fig. 2b). It has been argued that this crustacean has the tantalizing hexapod-like character of a missing second antenna[43]. Such Silurian fossils may generate important insights in the future.

Some version of a crustacean–hexapod relationship is now generally accepted, but the situation elsewhere on the euarthropod tree is more opaque. Several molecular studies support a surprising clade consisting of myriapods plus chelicerates, the so-called Paradoxopoda[30,44] (or Myriochelata[45]), which has little, if anything, to recommend it on morphological grounds[46]. Molecular support for Myriochelata is rather variable[25], suggesting that it may be a hard-to-resolve branching point. There is even weak support for Mandibulata being monophyletic when optimum out-groups, such as the priapulids, are used[47]. The early fossil record of myriapods is as unhelpful as that of the hexapods, possibly because both groups are terrestrial and so are less likely to be preserved.

The stem group of the chelicerates has been the subject of much recent interest[48], but it is not clear that any of the taxa proposed to lie in it, such as anomalocaridids and megacheirans (see the section 'Upper-stem-group euarthropods), belong there; a stem-group euarthropod placement may be more likely. Nevertheless, a group of taxa that have traditionally been considered to be stem-group chelicerates, including *Sidneyia*, and that have not been investigated extensively in the past 20 years, may yet offer important clues to early chelicerate evolution (Fig. 2c, d).

Arthropods emerged near the base of the Cambrian

Extant animal phyla are morphologically widely separated from other living clades and have long stem lineages that can, in principle, be reconstructed from the fossil record[49]. Several body fossils from the Ediacaran period, namely those of the genera *Spriggina* and *Parvancorina* and the 'soft-bodied trilobite' from Australia[50], have variously been described as arthropods[51] or even trilobites. All these assignments are highly questionable[52] for three reasons: first, they share no definite apomorphies with the arthropods; second, the forms that are superficially similar to the trilobites imply a considerable 'ghost lineage' of other arthropods, no trace of which is found in the Ediacaran record; and third, arthropod-like trace fossils are lacking until the Cambrian[53].

The earliest unequivocal records of arthropods are provided by trace fossils dating from shortly after the beginning of the Cambrian, considerably before the first undisputed body fossil[53]. The earliest form — simple scratches apparently made by arthropod limbs — belongs to the genus *Monomorphichnus* and is from the Early Cambrian in Newfoundland. These early traces are quickly joined by relatively large resting and burrowing traces from trace fossils assigned to the genera *Rusophycus* and *Cruziana*, which, although often thought to have been made by trilobites, may have resulted from the activity of any large animal with clawed limbs.

Euarthropods emerged from a diverse stem lineage

The body fossils of Cambrian stem-group arthropods have been sourced largely from the Chengjiang[54] and Sirius Passet[55] faunas of the Lower Cambrian, the Burgess Shale and other, related, sites of the Middle Cambrian[56], and the Orsten fauna[38] of the Upper Cambrian. Numerous lobopodians (including *Hallucigenia* species) that bear a superficial resemblance to the onychophorans have been described, but they cannot be demonstrated to have unequivocal onychophoran characteristics such as slime papillae[21]. As a result, they are best regarded as clustering around the last common ancestor of the onychophorans and euarthropods[57] (Fig. 3a, b).

In recent years, Cambrian lobopodian diversity has expanded to include several taxa — *Kerygmachela*[58], *Pambdelurion*[59] and *Megadictyon*[60] — that share some features with the first animals widely recognized to be stem-group euarthropods: the anomalocaridids (Fig. 3c) and the related *Opabinia*[61,62]. Anomalocaridids have clear euarthropod features such as sclerotized and articulating frontal appendages, large eyes on stalks and gut diverticula, but they lack other features, such as complete sclerotization of the cuticle.

Upper-stem-group euarthropods

Thirty years ago, little structure could be detected in the relationships of Cambrian arthropod fossils. However, in 1997, Xianguang Hou and Jan Bergström erected the Lamellipedia[63], a group of arthropods that includes trilobites and the related Trilobitoidea and is united by a biramous limb bearing lamellate setae on the outer branch (Fig. 3f). This clade resurrects some features of an earlier, often maligned, classification by Leif Størmer[64].

Two further clades or grades close to the euarthropods are now commonly recognized. The first clade is the Megacheira, or 'great appendage' arthropods[63,65]. This contains taxa such as *Leanchoilia* and *Yohoia* with rather uniform trunks and a single large anterior appendage (Fig. 3d). The second clade is as yet unnamed and includes relatives of the problematic genus *Fuxianhuia,* such as *Canadaspis* and *Perspicaris* species[66] (Fig. 3e).

Fuxianhuia has an anterior sclerite that has short stalked eyes, followed by a head shield with a pair of rather stout antenniform appendages[67]. Behind the antennae is a structure that may be part of the gut[68,69] but has also been described as a pair of 'subchelate' appendages, similar in form to those of the great appendage arthropods. The anterior sclerite has recently been identified in a wider range of taxa within the *Fuxianhuia* clade, notably *Branchiocaris*, which certainly bears a subchelate appendage[69]; such a sclerite is also present in some lamellipedians, such as *Helmetia* species, suggesting that lamellipedians may be paraphyletic[69].

These groupings — the lamellipedians, the megacheirans and the group containing *Fuxianhuia* — are now broadly accepted, but whether they are genuinely monophyletic or merely paraphyletic assemblages, as well as where they should be placed on the euarthropod and/or chelicerate and crustacean stem groups, remains unsettled. Much depends on the homology hypotheses made concerning the various anterior appendages and sclerites.

Most phylogenetic analyses have supported the placement of *Fuxianhuia* and its allies within the euarthropod stem group, but the lamellipedians, including the trilobites, have variously been placed in the upper stem group of the euarthropods[66] and in the stem groups of both mandibulates[70] and chelicerates[71] (Fig. 1). They may even be paraphyletic with (by implication) members on both the mandibulate and chelicerate lineages[69]. Likewise, the megacheirans have been placed either as a paraphyletic group slightly more derived than the anomalocaridids[66] or in the stem group to the chelicerates[8], possibly alongside the paraphyletic anomalocaridids[68] (Fig. 1). This troubling number of hypotheses stems from the agreed position of these fossils close to the euarthropod crown, which implies a scarcity of characters that are autapomorphic for modern clades.

The establishment of an outline phylogeny based on fossils and molecules, even with the caveats mentioned above, allows the evolution of crown-group euarthropod features to be traced through time. Next, we highlight one issue of particular current interest: the arthropod head problem.

The arthropod head problem

The problem of the composition of arthropod heads dates back more than a century. Although progress has been made, in particular into the homology of head segments of living taxa, one recurring question relates to the origin of the euarthropod labrum and the homology of the megacheiran great appendage to the appendages of living groups. If the great appendage is homologous to that of the anomalocaridids, a case can be made for tracing this appendage, via the anomalocaridid-like lobopodians *Pambdelurion* and *Kerygmachela*, to the antenna of the onychophorans[66], which has recently been shown to be protocerebral[72]. This would imply that neither the onychophoran antenna nor the great appendage is homologous to any overt appendage of living euarthropods, the most anterior appendage of which (the first antenna, or chelicera) is deutocerebral. However, as the great appendage seems to lie posterior to the antennae in at least some taxa, a rotation of the anterior end of the body is implied[66,70] (Fig. 4).

Conversely, denying the link between the anomalocaridids and *Kerygmachela*[65], or between the great appendage arthropods and the anomalocaridids[71], allows megacheirans to be placed in the stem group of the chelicerates, implying that the great appendage is homologous to the chelicerae and is therefore deutocerebral. Such a reconstruction could also imply that long, multi-articled antennae are a synapomorphy of the Mandibulata and that all taxa that have them, including the trilobites, belong in the mandibulate stem group[68].

Resolving this issue depends either on demonstrating convincing apomorphies that unite the anomalocaridids to *Pambdelurion* and *Kerygmachela,* placing them near the base of the arthropod tree, or on demonstrating a nested set of chelicerate features in purported stem-group chelicerates[48]. In addition, reconstruction of the stem-group record of the chelicerates will place important constraints on the morphological correlates, and the plausibility, of the Paradoxopoda hypothesis.

At the centre of the great appendage debate has been the labrum[70], a small structure found in front of the mouth in all extant euarthropods

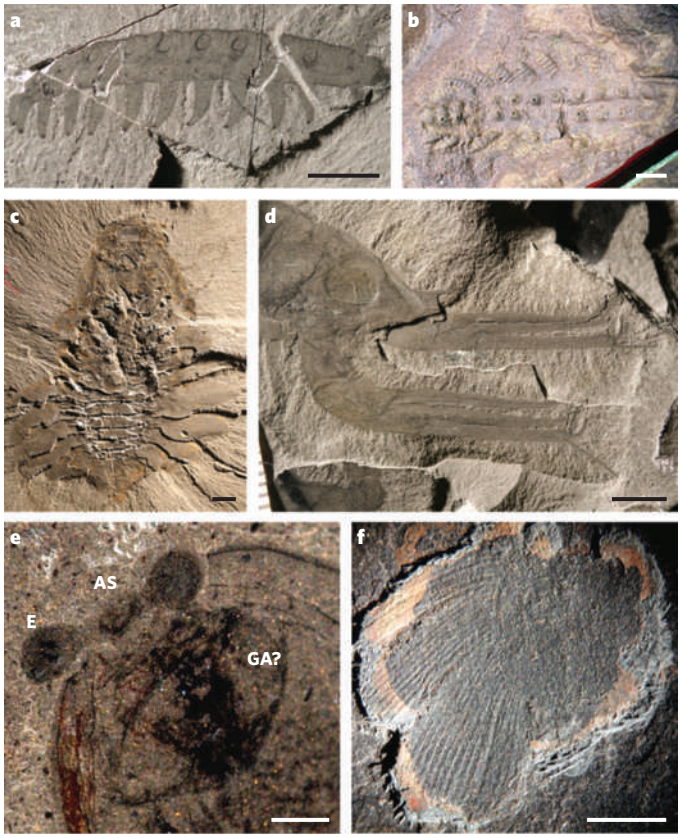


Figure 3 | Stem-group arthropods. a, *Aysheaia pedunculata* from the Burgess Shale (Middle Cambrian). Scale bar, 1 cm. **b**, *Xenusion auerswaldae* from Lower Cambrian sandstone erratic block, probably originally derived from Sweden. Scale bar, 1 cm. **c**, *Laggania cambria*, an anomalocaridid from the Burgess Shale. Scale bar, 1 cm. **d**, The great appendages of the megacheiran *Leanchoilia superlata* from the Burgess Shale. Scale bar, 0.5 cm. **e**, Head region of the probable *Fuxianhuia* relative *Perspicaris dictynna*, showing the eyes (E) attached to the 'anterior sclerite' (AS); note the large curved structures (GA?) behind this, which may correspond to the 'subchelate appendages' of *Fuxianhuia*. Scale bar, 1 mm. **f**, Outer limb branch of an undescribed Sirius Passet (Lower Cambrian of Greenland) lamellipedian arthropod, showing the characteristic flattened setae arranged along it. Scale bar, 0.5 cm. (Panels **a** and **c**–**e** courtesy of A. Daley.)

(except the pycnogonids, in which it may have been incorporated into the proboscis) and distinct from the plate-like hypostome of trilobites and other fossil taxa[73]. Although it is not universally accepted[73], most evidence points to the common origin of the labrum in all extant euarthropods as a protocerebrally derived structure that may be a modified appendage. Typically, the labrum develops at the anterior of the head as paired lobes (or at least its position is marked by paired expression domains of the appendage-patterning gene *Distal-less*) that fuse centrally and move posteriorly following the stomodeum (Fig. 4d, e). The anterior paired origin of the labrum and its possible appendiculate nature suggest homology with the anterior antennae of the onychophorans and the megacheiran great appendage[66] (Fig. 4). One test of this hypothesis would be a comparison of the molecular patterning of the onychophoran antenna and the euarthropod labrum[70], which the great appendage–labrum theory requires to be homologous structures.

Outstanding issues in arthropod evolution

Considerable progress has been made in the past few decades. The Ecdysozoa hypothesis has resulted in a shift of focus from the annelids as arthropod ancestors to the cycloneuralian worms, which will now receive much more attention. Similarly, dismantling the Atelocerata means that the shared features of myriapods and insects, such as tracheal breathing and Malpighian tubules for excretion, have been reinterpreted as convergent adaptations to life on land[74]. Gene-expression evidence

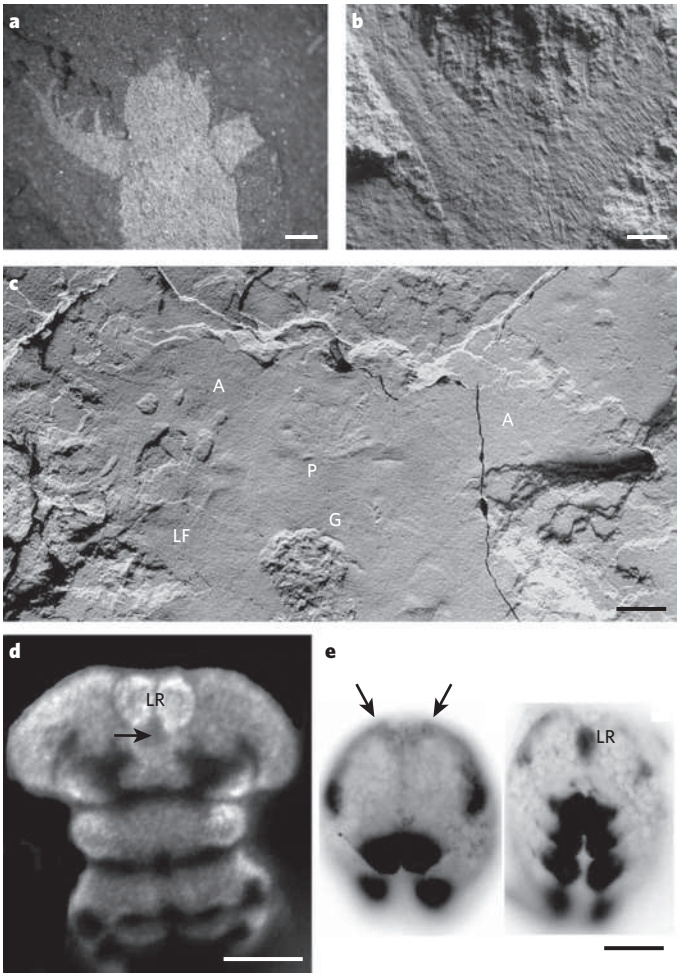


Figure 4 | Possible anterior-appendage homologies and mouth rotation in arthropods. a, The anterior mouth of *Aysheaia pedunculata*, flanked on either side by the spinose frontal appendages. Scale bar, 1 mm. **b**, Head region of *Kerygmachela kierkegaardi*, showing robust spinose appendages, which also flank the anterior mouth (not shown). Scale bar, 0.5 cm. **c**, Head region of *Pambdelurion whittingtoni*, showing the ventral mouth region with associated 'peytoia'-like (P) mouth structure and the two large and spinose frontal appendages (A) on either side. The gut (G) lies behind the peytoia mouthpart, together with the series of trunk lateral flaps (one marked LF). Scale bar, 1 cm. **d**, **e**, Labral development in extant arthropods. **d**, The paired anlagen of the labrum of the insect *Tribolium castaneum* before their posterior migration and merging to form the single adult labrum (labral anlagen marked with LR; stomodeal position indicated with arrow; counterstained with a probe that recognizes expression of the T-box gene *H15*). Scale bar, 50 µm. **e**, Progressive development of the labrum in the mite *Archegozetes longisetosus*. The labrum is marked by Distal-less-specific antibody staining (black) and develops from paired anterior regions (left, arrows), which then coalesce to form the adult labrum (right, LR). Scale bar, 25 µm. (Panel **d** courtesy of R. Janssen (Uppsala University, Sweden); panel **e** reproduced, with permission, from ref. 77.)

suggests that, in the case of insect and myriapod tracheae, the same pre-existing structures (gills attached to the base of appendages) may have been co-opted to this function in both cases, and this may also be the case for the air-breathing apparatus of terrestrial chelicerates[75]. Evo-devo research promises to supply fresh insights into all of these areas. Although it is clear that the myriapods are not the hexapod sister group, their true position remains enigmatic, and a solution must be found to this classic conflict between molecules and morphology.

Other key problems include the possible paraphyly of the cycloneuralian worms and the identity of the crustacean sister group to the insects, and indeed other hexapod taxa. Further developmental, morphological and molecular data are also required from putatively basal crustaceans, such as the remipedes and cephalocarids, for various mandibulate phylogeny hypotheses to be assessed.

One possible route to solving these phylogenetic issues is the provision of large new 'phylogenomic' data sets[76]. Entire genomes are being sequenced for diverse members of the Ecdysozoa, including a priapulid, a tardigrade, a horseshoe crab, a myriapod and an amphipod crustacean. As well as generating new data sets to help resolve large-scale ecdysozoan relationships, these will also provide insights into the ancestral ecdysozoan genome.

The relationships of the living arthropods may be resolved by phylogenomic approaches, but difficulties remain over the placement of fossil taxa close to the basal node of the crown euarthropods. These await new descriptions of fossils for resolution, although the statement of several distinct hypotheses now allows each of these to be tested using available material. Descriptions and systematization of stem-group arthropods from the fossil record have already allowed the origin of important arthropod features, such as limbs, sclerotization, head structures and even segmentation, to be traced deep into their history[62]. The debate engendered by the great appendage will no doubt rumble on, but it is conceivable that fossil embryos will be described that directly reveal the ontogeny of the relevant structures.

In the year of both the bicentenary of Darwin's birth and the centenary of the discovery of the Burgess Shale, it is clear that continued collaboration between molecular systematics, developmental biology and palaeontology will be required to resolve these outstanding issues in arthropod evolution. ■

1. Darwin, C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (John Murray, 1859).
2. Bowler, P. J. *Life's Splendid Drama: Evolutionary Biology and the Reconstruction of Life's Ancestry* (Chicago Univ. Press, 1996).
3. Ballard, J. W. O. *et al.* Evidence from 12S ribosomal-RNA sequences that onychophorans are modified arthropods. *Science* **258,** 1345–1348 (1992).
4. Telford, M. J., Bourlat, S. J., Economou, A., Papillon, D. & Rota-Stabelli, O. The evolution of the Ecdysozoa. *Phil. Trans. R. Soc. Lond. B* **363,** 1529–1537 (2008).
5. Dunn, C. W. *et al.* Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* **452,** 745–749 (2008).
This paper provides the most recent analysis of the relationships of the Metazoa, using a broadly sampled phylogenomic-scale data set.
6. Dzik, J. & Krumbiegel, G. The oldest 'onychophoran' *Xenusion*: a link connecting phyla? *Lethaia* **22,** 169–182 (1989).
7. Eernisse, D. J., Albert, J. S. & Anderson, F. E. Annelida and Arthropoda are not sister taxa — a phylogenetic analysis of spiralian metazoan morphology. *Syst. Biol.* **41,** 305–330 (1992).
8. Aguinaldo, A. M. A. *et al.* Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387,** 489–493 (1997).
This classic paper provided the first strong evidence for the clade Ecdysozoa.
9. Wagele, J. W., Erikson, T., Lockhart, P. & Misof, B. The Ecdysozoa: artifact or monophylum? *J. Zoolog. Syst. Evol. Res.* **37,** 211–223 (1999).
10. Philip, G. K., Creevey, C. J. & McInerney, J. O. The Opisthokonta and the Ecdysozoa may not be clades: stronger support for the grouping of plant and animal than for animal and fungi and stronger support for the Coelomata than Ecdysozoa. *Mol. Biol. Evol.* **22,** 1175–1184 (2005).
11. Rogozin, I. B., Thomson, K., Csueroes, M., Carmel, L. & Koonin, E. V. Homoplasy in genome-wide analysis of rare amino acid replacements: the molecular-evolutionary basis for Vavilov's law of homologous series. *Biol. Direct* **3,** 7 (2008).
12. Rogozin, I. B., Wolf, Y. I., Carmel, L. & Koonin, E. V. Analysis of rare amino acid replacements supports the Coelomata clade. *Mol. Biol. Evol.* **24,** 2594–2597 (2007).
13. Rogozin, I. B., Wolf, Y. I., Carmel, L. & Koonin, E. V. Ecdysozoan clade rejected by genome-wide analysis of rare amino acid replacements. *Mol. Biol. Evol.* **24,** 1080–1090 (2007).
14. Zheng, J., Rogozin, I. B., Koonin, E. V. & Przytycka, T. M. Support for the Coelomata clade of animals from a rigorous analysis of the pattern of intron conservation. *Mol. Biol. Evol.* **24,** 2583–2592 (2007).
15. Copley, R. R., Aloy, P., Russell, R. B. & Telford, M. J. Systematic searches for molecular synapomorphies in model metazoan genomes give some support for Ecdysozoa after accounting for the idiosyncrasies of *Caenorhabditis elegans*. *Evol. Dev.* **6,** 164–169 (2004).
16. Philippe, H., Lartillot, N. & Brinkmann, H. Multigene analyses of bilaterian animals corroborate the monophyly of Ecdysozoa, Lophotrochozoa, and Protostomia. *Mol. Biol. Evol.* **22,** 1246–1253 (2005).
17. Webster, B. L. *et al.* Mitogenomics and phylogenomics reveal priapulid worms as extant models of the ancestral Ecdysozoan. *Evol. Dev.* **8,** 502–510 (2006).
18. Irimia, M., Maeso, I., Penny, D., Garcia-Fernàndez, J. & Roy, S. Rare coding sequence changes are consistent with Ecdysozoa, not Coelomata. *Mol. Biol. Evol.* **24,** 1604–1607 (2007).
19. Papillon, D., Perez, Y., Caubit, X. & Le Parco, Y. Identification of chaetognaths as protostomes is supported by the analysis of their mitochondrial genome. *Mol. Biol. Evol.* **21,** 2122–2129 (2004).
20. Schmidt-Rhaesa, A., Bartolomaeus, T., Lemburg, C., Ehlers, U. & Garey, J. R. The position of the Arthropoda in the phylogenetic system. *J. Morphol.* **238,** 263–285 (1998).
21. Budd, G. E. The morphology and phylogenetic significance of *Kerygmachela kierkegaardi* Budd (Buen Formation, Lower Cambrian, N Greenland). *Trans. R. Soc. Edinb.* **89,** 249–290 (1999).

22. Eriksson, B. J. & Budd, G. E. Onychophoran cephalic nerves and their bearing on our understanding of head segmentation and stem-group evolution of Arthropoda. *Arthropod Struct. Dev.* **29,** 197–209 (2000).
23. Nielsen, C. *Animal Evolution: Interrelationships of the Animal Phyla* (Oxford Univ. Press, 2001).
24. Mallatt, J. & Giribet, G. Further use of nearly complete, 28S and 18S rRNA genes to classify Ecdysozoa: 37 more arthropods and a kinorhynch. *Mol. Phylogenet. Evol.* **40,** 772–794 (2006).
25. Bourlat, S. J., Nielsen, C., Economou, A. D. & Telford, M. J. Testing the new animal phylogeny: a phylum level analysis of the animal kingdom. *Mol. Phylogenet. Evol.* **49,** 23–31 (2008).
26. Manton, S. M. *The Arthropoda: Habits, Functional Morphology and Evolution* (Clarendon, 1977).
27. Fortey, R. A., Briggs, D. E. G. & Wills, M. A. The Cambrian evolutionary 'explosion': decoupling cladogenesis from morphological disparity. *Biol. J. Linn. Soc.* **57,** 13–33 (1996).
28. Boore, J. L., Lavrov, D. V. & Brown, W. M. Gene translocation links insects and crustaceans. *Nature* **392,** 667–668 (1998).
This paper details a rare genomic change supporting Pancrustacea rather than Atelocerata.
29. Giribet, G., Edgecombe, G. D. & Wheeler, W. C. Arthropod phylogeny based on eight molecular loci and morphology. *Nature* **413,** 157–161 (2001).
30. Mallatt, J. M., Garey, J. R. & Shultz, J. W. Ecdysozoan phylogeny and Bayesian inference: first use of nearly complete 28S and 18S rRNA gene sequences to classify the arthropods and their kin. *Mol. Phylogenet. Evol.* **31,** 178–191 (2004).
31. Richter, S. The Tetraconata concept: hexapod-crustacean relationships and the phylogeny of Crustacea. *Org. Divers. Evol.* **2,** 217–237 (2002).
32. Ungerer, P. & Scholtz, G. Filling the gap between identified neuroblasts and neurons in crustaceans adds new support for Tetraconata. *Proc. R. Soc. B* **275,** 369–376 (2008).
This paper presents data typical of the high-quality morphological techniques now being used to address arthropod phylogeny.
33. Harzsch, S. & Hafner, G. Evolution of eye development in arthropods: phylogenetic aspects. *Arthropod Struct. Dev.* **35,** 319–340 (2006).
34. Fanenbruck, M. & Harzsch, S. A brain atlas of *Godzilliognomus frondosus* Yager, 1989 (Remipedia, Godzilliidae) and comparison with the brain of *Speleonectes tulumensis* Yager, 1987 (Remipedia, Speleonectidae): implications for arthropod relationships. *Arthropod Struct. Dev.* **34,** 343–378 (2005).
35. Sanders, H. L. The Cephalocarida and crustacean phylogeny. *Syst. Zool.* **6,** 112–128 (1957).
36. Cook, C. E., Yue, Q. Y. & Akam, M. Mitochondrial genomes suggest that hexapods and crustaceans are mutually paraphyletic. *Proc. R. Soc. B* **272,** 1295–1304 (2005).
37. Carapelli, A., Liò, P., Nardi, F., van der Wath, E. & Frati, F. Phylogenetic analysis of mitochondrial protein coding genes confirms the reciprocal paraphyly of Hexapoda and Crustacea. *BMC Evol. Biol.* **7** (suppl. 2), S8 (2007).
38. Siveter, D. J., Williams, M. & Waloszek, D. A phosphatocopid crustacean with appendages from the Lower Cambrian. *Science* **293,** 479–481 (2001).
39. Zhang, X. G., Siveter, D. J., Waloszek, D. & Maas, A. An epipodite-bearing crown-group crustacean from the Lower Cambrian. *Nature* **449,** 595–598 (2007).
40. Siveter, D. J., Sutton, M. D., Briggs, D. E. G. & Siveter, D. J. A new probable stem lineage crustacean with three-dimensionally preserved soft parts from the Herefordshire (Silurian) Lagerstatte, UK. *Proc. R. Soc. B* **274,** 2099–2107 (2007).
This paper provides valuable new data on a crustacean-like taxon considerably younger than the Cambrian faunas.
41. Wilson, H. M. Juliformian millipedes from the Lower Devonian of Euramerica: implications for the timing of millipede cladogenesis in the Paleozoic. *J. Paleontol.* **80,** 638–649 (2006).
42. Fayers, S. R. & Trewin, N. H. A hexapod from the Early Devonian Windyfield chert, Rhynie, Scotland. *Palaeontology* **48,** 1117–1130 (2005).
43. Boxshall, G. A. Crustacean classification: on-going controversies and unresolved problems. *Zootaxa* **1668,** 313–325 (2007).
44. Friedrich, M. & Tautz, D. Ribosomal DNA phylogeny of the major extant arthropod classes and the evolution of myriapods. *Nature* **376,** 165–167 (1995).
45. Pisani, D., Poling, L., Lyons-Weiler, M. & Hedges, S. The colonization of land by animals: molecular phylogeny and divergence times among arthropods. *BMC Biol.* **2,** 1 (2004).
46. Edgecombe, G. D. Morphological data, extant Myriapoda, and the myriapod stem-group. *Contrib. Zool.* **73,** 207–252 (2004).
47. Rota-Stabelli, O. & Telford, M. J. A multi criterion approach for the selection of optimal outgroups in phylogeny: recovering some support for Mandibulata over Myriochelata using mitogenomics. *Mol. Phylogenet. Evol.* **48,** 103–111 (2008).
48. Dunlop, J. A. New ideas about the euchelicerate stem-lineage. *Acta Zool. Bulg.* (Suppl. 1) 9–23 (2005).
49. Budd, G. E. The Cambrian fossil record and the origin of the phyla. *Integr. Comp. Biol.* **43,** 157–165 (2003).
50. Gehling, J. G. The case for Ediacaran fossil roots to the metazoan tree. *Mem. Geol. Soc. India* **20,** 181–223 (1991).
51. Waggoner, B. M. Phylogenetic hypotheses of the relationships of arthropods to Precambrian and Cambrian problematic fossil taxa. *Syst. Biol.* **45,** 190–222 (1996).
52. Ivantsov, A. Y. Vendia and other Precambrian 'arthropods.' *Paleontol. J.* **35,** 335–343 (2001).
53. Jensen, S. The Proterozoic and earliest Cambrian trace fossil record; patterns, problems and perspectives. *Integr. Comp. Biol.* **43,** 219–228 (2003).
54. Hou, X.-G., Aldridge, R. J., Bergström, J., Siveter, D. J. & Feng, X.-H. *The Cambrian Fossils of Chengjiang, China* (Blackwell Science, 2004).
55. Conway Morris, S., Peel, J. S., Higgins, A. K., Soper, N. J. & Davis, N. C. A Burgess Shale-like fauna from the Lower Cambrian of North Greenland. *Nature* **326,** 181–183 (1987).
56. Briggs, D. E. G. & Collins, D. The arthropod *Alalcomenaeus cambricus* Simonetta, from the Middle Cambrian Burgess Shale of British Columbia. *Palaeontology* **42,** 953–977 (1999).
57. Whittington, H. B. The lobopod animal *Aysheaia pedunculata* Walcott, Middle Cambrian, Burgess Shale, British Columbia. *Phil. Trans. R. Soc. Lond. B* **284,** 165–197 (1978).
58. Budd, G. A Cambrian gilled lobopod from Greenland. *Nature* **364,** 709–711 (1993).
59. Budd, G. E. Arthropod body-plan evolution in the Cambrian with an example from anomalocaridid muscle. *Lethaia* **31,** 197–210 (1998).
60. Liu, J. N., Shu, D. G., Han, J., Zhang, Z. F. & Zhang, X. L. Morpho-anatomy of the lobopod *Magadictyon* cf. *haikouensis* from the Early Cambrian Chengjiang Lagerstatte, South China. *Acta Zool.* **88,** 279–288 (2007).
61. Bergstrom, J. *Opabinia* and *Anomalocaris*, unique Cambrian arthropods. *Lethaia* **19**, 241–246 (1986).
62. Budd, G. E. The morphology of *Opabinia regalis* and the reconstruction of the arthropod stem-group. *Lethaia* **29,** 1–14 (1996).
63. Hou, X.-G. & Bergström, J. *Fossils and Strata* Vol. 45 (Wiley, 1997).
64. Størmer, L. On the relationship and phylogeny of fossil and recent Arachnomorpha. A comparative study in Arachnida, Xiphosura, Eurypterida, Trilobita and other fossil Arthropoda. *Skr. Norske Vidensk-Akad.* **5,** 1–158 (1944).
65. Chen, J. Y., Waloszek, D. & Maas, A. A new 'great-appendage' arthropod from the Lower Cambrian of China and homology of chelicerate chelicerae and raptorial antero-ventral appendages. *Lethaia* **37,** 3–20 (2004).
66. Budd, G. E. A palaeontological solution to the arthropod head problem. *Nature* **417,** 271–275 (2002).
67. Chen, J. Y., Edgecombe, G. D., Ramskold, L. & Zhou, G. Q. Head segmentation in Early Cambrian *Fuxianhuia* — implications for arthropod evolution. *Science* **268,** 1339–1343 (1995).
68. Waloszek, D., Chen, J. Y., Maas, A. & Wang, X. Q. Early Cambrian arthropods — new insights into arthropod head and structural evolution. *Arthropod Struct. Dev.* **34,** 189–205 (2005).
69. Budd, G. E. Head structure in upper stem-group euarthropods. *Palaeontology* **51,** 561–573 (2008).
70. Scholtz, G. & Edgecombe, G. D. The evolution of arthropod heads: reconciling morphological, developmental and palaeontological evidence. *Dev. Genes Evol.* **216,** 395–415 (2006).
71. Cotton, T. J. & Braddy, S. J. The phylogeny of arachnomorph arthropods and the origin of the Chelicerata. *Trans. R. Soc. Edinb.* **94,** 169–193 (2004).
72. Eriksson, B. J., Tait, N. N. & Budd, G. E. Head development in the onychophoran *Euperipatoides kanangrensis* with particular reference to the central nervous system. *J. Morphol.* **255,** 1–23 (2003).
73. Walossek, D. & Muller, K. J. Upper Cambrian stem-lineage crustaceans and their bearing upon the monophyletic origin of Crustacea and the position of *Agnostus*. *Lethaia* **23,** 409–427 (1990).
74. Telford, M. J. & Thomas, R. H. Demise of the Atelocerata? *Nature* **376,** 123–124 (1995).
75. Damen, W. G. M., Saridaki, T. & Averof, M. Diverse adaptations of an ancestral gill: a common evolutionary origin for wings, breathing organs, and spinnerets. *Curr. Biol.* **12,** 1711–1716 (2002).
The fascinating gene-expression data in this paper support the likely homology of insect wings and tracheae with crustacean gills.
76. Philippe, H. & Telford, M. J. Large-scale sequencing and the new animal phylogeny. *Trends Ecol. Evol.* **21,** 614–620 (2006).
77. Thomas, R. H. & Telford, M. J. Appendage development in embryos of the oribatid mite *Archegozetes longisetosus* (Acari, Oribatei, Trhypochthoniidae). *Acta Zool.* **80,** 193–200 (1999).

Acknowledgements We thank A. Daley for assistance with photography and R. Copley for critically reading an earlier draft. Our research is supported by the Marie Curie Research Training Network ZOONET (grant number MRTN-CT-2004-005624).

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence should be addressed to the authors (graham.budd@pal.uu.se; m.telford@ucl.ac.uk).

NATURE|Vol 457|12 February 2009|doi:10.1038/nature07891

Deep homology and the origins of evolutionary novelty

Neil Shubin[1,2], Cliff Tabin[3] & Sean Carroll[4]

Do new anatomical structures arise *de novo*, or do they evolve from pre-existing structures? Advances in developmental genetics, palaeontology and evolutionary developmental biology have recently shed light on the origins of some of the structures that most intrigued Charles Darwin, including animal eyes, tetrapod limbs and giant beetle horns. In each case, structures arose by the modification of pre-existing genetic regulatory circuits established in early metazoans. The deep homology of generative processes and cell-type specification mechanisms in animal development has provided the foundation for the independent evolution of a great variety of structures.

It is not possible to identify what is new in evolution without understanding the old. This is a reflection of the way evolution works, with some novelties being traceable as modifications of primitive conditions and others having origins that are much less obvious. As a result, the problems of novelty and homology have been deeply intertwined for the past century and a half. It is here, at the interface between these two great concepts of evolutionary biology, where fresh data from developmental biology have had an extraordinary impact. One of the most important, and entirely unanticipated, insights of the past 15 years was the recognition of an ancient similarity of patterning mechanisms in diverse organisms, often among structures not thought to be homologous on morphological or phylogenetic grounds. In 1997, prompted by the remarkable extent of similarities in genetic regulation between organs as different as fly wings and tetrapod limbs, we suggested the term 'deep homology'[1] to describe the sharing of the genetic regulatory apparatus that is used to build morphologically and phylogenetically disparate animal features[1,2].

Homology, as classically defined, refers to a historical continuity in which morphological features in related species are similar in pattern or form because they evolved from a corresponding structure in a common ancestor. Deep homology also implies a historical continuity, but in this case the continuity may not be so evident in particular morphologies; it lies in the complex regulatory circuitry inherited from a common ancestor. In some instances, recognition of deep homologies can help in the identification of cryptic classical homologies, when morphological data alone are inadequate to make the case for homology. For example, the photoreceptors present in various extant clades would not be recognized as homologous without the observation of common underlying genetic cassettes (discussed in the next section).

Deep homology, however, can also be found in contexts in which structures are not homologous in the classical sense. As we explored in 1997, appendages in vertebrates, arthropods and other bilaterians evolved independently, but their derivation was dependent on regulatory networks present in a common 'urbilaterian' ancestor. Most strikingly, the genetic regulatory cascade comprising a key transcription factor and downstream effector genes eliciting outgrowth (such as the *Drosophila melanogaster* gene *Distal-less* or its mouse homologue *Dlx*) seems to have been present in such a common ancestor and has been repeatedly used to control outgrowth formation in the protostome and deuterostome lineages[3]. Moreover, a series of deep homologies exist in the genetic systems used to pattern the appendages of vertebrates and arthropods, many of which have come to light since our original paper was written (for example, proximal-appendage specification by *homothorax* in *D. melanogaster* or its homologue *Meis1* in mice[4,5]). The similarities are much more than the use of a common genetic tool kit of genes: they involve the use of genes and regulatory circuits that have previously evolved complex roles in an ancestral organism.

Deep homology is important for the generation of novelties because ancient regulatory circuits provide a substrate from which novel structures can develop. In this Review, we explore three of Charles Darwin's exemplars of evolution: animal eyes, tetrapod limbs and the giant horns of beetles. New data from studies of these features are offering surprising twists on classic examples of evolution. And, together, these examples illustrate how deep homology enables researchers to understand the generation of novelty in cases in which fossils are not informative; to make predictions about morphological transformation that can be tested by experimental and expeditionary work; and to see the extent to which common genetic mechanisms are used to generate diverse adaptations and can lead to the parallel evolution of novelties.

What eyes tell us

Darwin noted in *On the Origin of Species*[6] that "Amongst existing Vertebrata, we find but a small amount of gradation in the structure of the eye, and from fossil species we can learn nothing on this head. In this great class we should probably have to descend far beneath the lowest known fossiliferous stratum to discover the earlier stages, by which the eye has been perfected."

Unless an animal lives in complete darkness, there are tremendous advantages to being able to sense light and have high-acuity vision. Accordingly, complex eyes are present in a wide variety of taxa, from jellyfish and molluscs to insects and vertebrates. But how and when did they develop?

The eyes in widely divergent organisms are of such different structures, and develop in such distinct embryological contexts, that biologists

[1]Department of Organismal Biology and Anatomy, University of Chicago, 1027 East 57th Street, Chicago, Illinois 60637, USA. [2]The Field Museum of Natural History, 1400 South Lake Shore Drive, Chicago, Illinois 60605, USA. [3]Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, NRB 360k, Boston, Massachusetts 02115, USA. [4]Howard Hughes Medical Institute, University of Wisconsin-Madison, 1525 Linden Drive, Madison, Wisconsin 53706, USA.

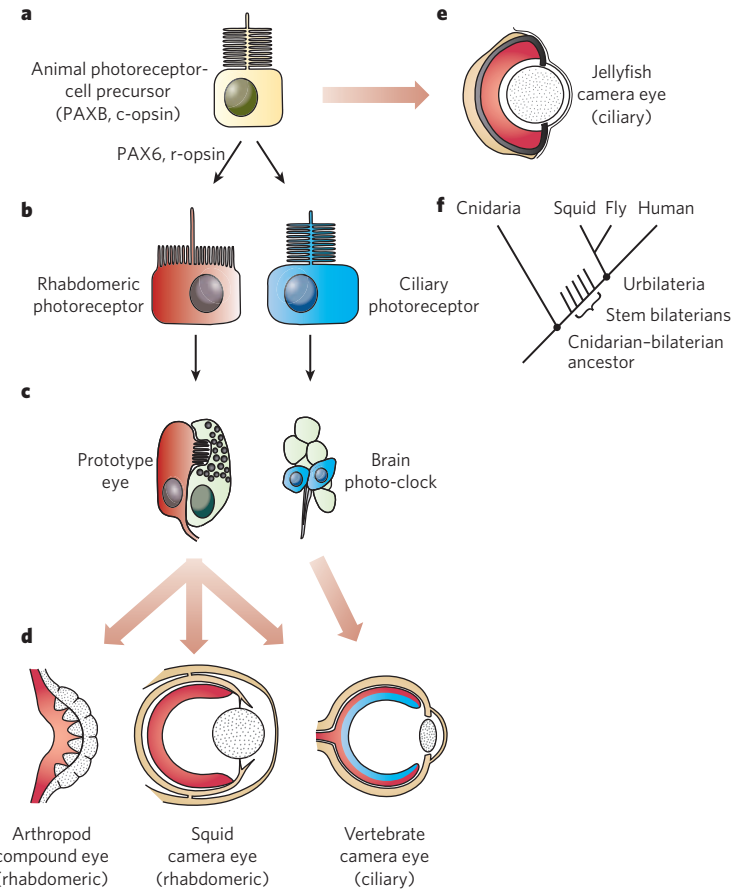


Figure 1 | Deep homology of eye development and the parallel evolution of animal eyes. A route for the evolution of photoreceptor cell types and different forms of eyes. **a**, The cnidarian–bilaterian ancestor had photoreceptors that expressed c-opsin and PAXB. **b**, Rhabdomeric photoreceptors, r-opsins and PAX6 evolved in ancestral-stem bilaterians, after the split between the cnidarian and bilaterian lineages. **c**, The last common ancestor of all bilaterians (Urbilateria) probably had two types of light-sensing organ: a prototypical eye and a brain photo-clock, which are both found in the annelid *Platynereis dumerilii*. **d**, The photoreceptor types established in the Urbilateria were then incorporated in different ways in the parallel evolution of different eyes in various phyla. Rhabdomeric photoreceptors were the foundation for the evolution of compound and camera-type eyes in arthropods and molluscs, respectively. Both types of photoreceptor were incorporated into the vertebrate eye, with ciliary receptors carrying out phototransduction and rhabdomeric receptors being transformed into ganglion cells and functioning in image processing. Pigment cells are not shown. **e**, The ciliary camera-type eyes of box jellyfish are also proposed to have evolved in parallel in the cnidarian lineage. **f**, Cladogram depicting the evolutionary relationships of the taxa shown in **a**–**e**. (Panels **a**–**e** courtesy of L. Olds (University of Wisconsin, Madison); panel **f** courtesy of K. Monoyios (University of Chicago, Illinois).)

have historically inferred that eyes had evolved independently dozens of times[7]. But deciding where homology ends and novelty begins is not always straightforward, particularly for structures that lack a fossil record, such as eyes. Over the past 15 years, many insights into the evolution of eyes have come from descending beneath the visible diversity of animal eyes into the genetic machinery that controls their development.

The unexpected finding that the homologous transcription factors Eyeless and PAX6 have crucial roles in the formation of the eyes of *D. melanogaster* and vertebrates was the first indication that the markedly different eyes of long-diverged phyla had more in common than was previously thought[4]. This discovery spurred comparisons of the detailed genetic circuitry underlying eye formation in diverse animals. It is now known that a small set of transcription factors, including those encoded by members of the *eyeless, atonal* and *eyes absent* gene families in *D. melanogaster* and their homologues in vertebrates, are widely used in the specification and formation of various types of animal eye (see refs 8 and 9 for reviews). Moreover, in all light-sensing organs in animals, the ability to detect light depends on a cascade involving opsin proteins. This observation led to the view that all modern variations of light sensing in bilaterians can be traced to the existence of photosensitive cells in a common ancestor with PAX6 and other transcription factors at the top of a genetic regulatory pathway leading to opsin production. This is a textbook example of deep homology[1,2]: morphologically disparate organs whose formation (and evolution) depends on homologous genetic regulatory circuits.

The deep homology of regulatory mechanisms suggests a number of different hypotheses for evolutionary transformation[10,11]. Are all eyes homologous in the classical sense of being diverse forms of the same structure? Have different eyes instead evolved entirely independently and just happened to use similar genetic components? Or have different eyes evolved in parallel as elaborations of structures and genetic regulatory mechanisms present in common ancestors? Weighing up these possibilities requires a deeper examination of eye development and a broader survey of taxa. It turns out that far more striking developmental similarities exist between diverse taxa than would ever have been expected. However, there are also many cases in which, on close inspection of the details, the similarities break down.

One well-established argument for the independent evolution of eyes in arthropods, squid and vertebrates concerned their photoreceptor composition and light-transduction mechanisms. All animal eyes are composed of photoreceptors located adjacent to cells that produce light-shielding dark pigment. The photoreceptors of different taxa are of two main types, rhabdomeric and ciliary (Fig. 1b), which have distinct phototransduction signalling cascades. The eyes of insects and other invertebrates rely on rhabdomeric photoreceptors and a phospholipase-C-based cascade, whereas vertebrates use ciliary photoreceptors and a phosphodiesterase-based cascade. How can the fact that similar camera-type eyes of vertebrates and squids rely on different kinds of photoreceptor and transduction process be explained, other than by asserting completely separate origins?

It is clear that these photoreceptors were not separate inventions in each respective lineage. In polychaete annelids, both photoreceptor types are found, with the ciliary-type opsin (c-opsin) expressed in the brain and rhabdomeric opsin (r-opsin) expressed in the eyes of the same animal[12]. The best explanation for the phylogenetic distribution of photoreceptor types is that both cell types coexisted in the common bilaterian ancestor of vertebrates and invertebrates (Fig. 1c), and each lineage then used different cell types for light detection in their visual systems. Both types of photoreceptor seem to have been incorporated into the evolving vertebrate eye, with the rhabdomeric photoreceptor cells being transformed into ganglion cells and having a key role in image processing[12] (Fig. 1d). So one fundamental element of the deep homology of eye development lies in the specification of cell types.

The deep homologies are not limited to the photoreceptors. The target interneurons specifically involved in processing the visual signal in the vertebrate retina and the *D. melanogaster* optic lobe are not only functionally analogous but are specified by homologous transcription factors (CHX10 and VSX2). Moreover, both cell types project to visual centres of the brain, and both sequentially express genes encoding the

same two transcription factors: first, MATH5 or its homologue Atonal, and then BRN3B or ACJ6 (in mice and *D. melanogaster*, respectively)[13]. These findings imply that these different bilaterian eyes are related, not just in the homology of cell types but in the deep homology of cellular circuitry for interpreting visual input, which predates the independent evolution of the modern eyes in the two lineages.

How deep do the cell-type homologies extend in the phylogenetic tree of metazoans? The cnidarians, the likely sister group to the bilaterians, include the box jellyfish, which has a camera-type eye with a cornea, a lens and a retina. These animals not only have retinas with c-opsin, typical of vertebrate-type eyes[14,15], but also express the full set of orthologous genes used in vertebrate phototransduction[16]. Furthermore, the pigment function is combined in jellyfish photoreceptors that use the same melanogenic pathway seen in vertebrates, involving, for example, *mitf* and a homologue of *Oca2* (ref. 16).

The cnidarian photoreceptors reveal that ciliary photoreceptors, c-opsin and the phosphodiesterase signal-transduction cascade predate the divergence between cnidarians and bilaterians and, because r-opsins have not been detected in cnidarians, it is probable that r-opsins are a bilaterian innovation[17,18]. The anatomical resemblances between cnidarian and vertebrate camera-type eyes perhaps reflect conservation of a camera-type eye 'program' and consequent parallel evolution of eyes in the two phyla (Fig. 1a). If a ciliary-type camera-style eye was assembled in a cnidarian–bilaterian ancestor, it must have been lost in other bilaterians except for a vertebrate ancestor. Given the advantages of vision, it is difficult to imagine why such structure would be lost before being reinvented using rhabdomeric photoreceptors.

Taken together, the available data suggest that the eyes of jellyfish, squid, arthropods and vertebrates are not the product of rampant convergent evolution, as once thought, but of parallel evolution that is based on a shared history of generative mechanisms and cell types — deep homologies — established early in animal evolution.

Tetrapod limbs and fish fins

In *On the Origin of Species*, Darwin wrote of the similarities of tetrapod limbs: "What can be more curious than that the hand of a man, formed for grasping, that of a mole for digging, the leg of the horse, the paddle of the porpoise, and the wing of the bat, should all be constructed on the same pattern, and should include similar bones, in the same relative positions?"

The origin of limbs in the Devonian period allowed the invasion of land and the later evolution of vertebrates that could fly, dig, run, hop and climb. Consequently, tetrapod limbs are classic examples of evolutionary novelties[19,20]. But they are also a prime example of homology: all tetrapod limbs have similar bone morphology and development, and this can be traced back to the limbs of Devonian vertebrates. As was clear to Darwin, this homology is immediately apparent in the detailed similarities of morphology and development of the bones in the limbs of all tetrapods. More controversial, however, have been attempts to compare these bones with those in the paired fins of fish. It is here where deep genetic homologies and the discovery of fossils conspire to offer fresh insights.

The most striking differences between limbs and fins lie in the distal region of the appendage, as the pattern of the proximal bones is essentially identical among lobe-finned fish and tetrapods[19,20]. Much of the surface area of fins is supported by dermal rays — bones that are completely absent in limbs. In place of these rays, limbs have a set of endochondral elements — the digits and wrist or ankle bones — that look unlike, and function differently from, fin radials. Importantly, these bones contain a series of characteristic joints that allow flexion and extension, particularly in those that enable a 'palm' area to lie flush with the ground[21]. Tetrapods, then, have traded a dermal skeleton in the distal appendage for a complex endochondral one, and their evolved joints and bones allow the distal appendage to support the weight of the body.

Genetic discoveries in the 1990s reinforced a classical view that digits, wrists and ankles have no direct correlate in fins[1,22–25]. Developmental studies of *Hoxd9*, *Hoxd10*, *Hoxd11*, *Hoxd12* and *Hoxd13* gene expression in tetrapod limbs revealed a discrete phase of expression directly associated with digital specification[22,23]. There are four important features of this late-phase expression: first, expression occurs in a distal segment of the limb, the 'paddle', that does not overlap with more proximal zones; second, this expression occurs while the digits and mesopodial bones are being specified; third, the domains of expression of the 5′ *Hoxd* genes display 'reverse colinearity', such that *Hoxd13* expression (for example) extends more anteriorly than the expression domains of more 3′ genes; and fourth, late-phase expression is regulated independently from early-phase expression by a separate enhancer that drives distal expression[26]. This distinct expression pattern of the *Hoxd* genes has been referred to as 'phase 2 expression'. Phase 2 expression, with all the characteristics described above, was unknown in fish fins in the late 1990s, as the zebrafish (*Danio rerio*) was reported to have only a single phase of expression[23]. The fossil record also seemed to support this conclusion, as the sister group to limbed vertebrates (panderichthyids) seemed to lack any bones comparable to digits[1,19,20]. Data from both types of study supported the idea that the origin of novelty at the morphological and functional levels would have happened in parallel to that at the genetic level.

Just as for eyes, comparative data refined the hypothesis of novelty (Fig. 2). Work on a variety of non-model vertebrates has revealed that a late phase of *Hoxd* expression exists in the distal fin bud. Indeed, it is now known to be a general feature of gnathostomes, having been discovered in basal actinopterygians[27], lungfish[28], zebrafish[29] and a chondrichthyan[30]. Although the details of late-phase patterns of expression vary between these taxa, some but not all aspects of the tetrapod phase 2 *Hox* pattern are present in fins. The most notable difference is that late-phase expression in the fins of osteichthyans (paddlefish, zebrafish and lungfish) spatially overlaps with earlier phases of expression, whereas in tetrapods the phases of expression are segregated proximodistally such that phase 2, but not phase 1, expression is found in the autopod. Despite these differences, basal actinopterygians and lungfish have broad zones of expression that, like tetrapods, exhibit reverse colinearity.

The key question is whether these late phases of expression in tetrapod limbs and fish fins reflect the same process. There is, as yet, no evidence of independent regulation of the form described in tetrapods. Lacking this evidence, there are two tenable hypotheses: either fish have phase 2 expression homologous to that which specifies the digits in tetrapods, or the late-phase expression observed in fins is a temporal and spatial extension of a highly conserved early phase with unique dynamic properties in the fish. The two cases suggest different molecular scenarios for the origin of the autopod. If fish have true phase 2 *Hox* gene expression, the main difference between tetrapod limbs and fish fins lies not in the origin of late-phase *Hox* gene expression but in changes in the timing of this expression and/or in changes in genes acting downstream of the *Hox* genes that must have shaped the tetrapod form. Alternatively, if fish lack true phase 2 *Hox* gene expression, the evolution of the phase 2 regulatory module would be a unique and perhaps defining autopod invention of the tetrapod lineage. In either case, the evolutionary process reveals the impact of deep homology: different kinds of appendage arose by modifications to an ancient and conserved developmental system. Even if the morphological structures (fin rays and digits) are not homologous, there is deeper homology in the network of *Hox* genes and their targets within the limb or fin field, as well as in phase 2 regulation — if indeed it predated the evolution of the autopod.

From experiments to expeditions

If components of the genetic machinery that builds the skeleton of limbs were present in fins, the closest fossil relatives of Devonian tetrapods may indeed have had equivalents of wrist and ankle bones or digits. In developmental genetics, the database is expanded by carrying out experiments on different taxa. In palaeontology, knowledge is expanded by discovering new fossil forms. The origin of limbs is a problem that is ripe for analysis. When cladograms of basal limbed vertebrates and their finned relatives are placed on a stratigraphic column, it becomes apparent that new information on the fin-to-limb transformation is most likely to come from expeditionary work in the

Frasnian and Givetian stages (385 million to 375 million years ago) of the Devonian. For more than 50 years, it has been known that some lobe-finned fish have homologues of two wrist bones, the intermedium and ulnare. But what was not known until recently is what the endoskeleton distal to these bones looked like in the closest relatives of tetrapods and how it functioned.

Targeting the Frasnian rocks of Arctic Canada for fossil vertebrates ultimately led to the discovery of *Tiktaalik roseae,* the closest finned sarcopterygian to limbed forms[21,31]. *T. roseae* is highly informative because most of the skeleton is exceptionally well preserved. In addition, because its appendages are known from multiple specimens, pectoral fins have been prepared both in articulation and as isolated elements preserved in three dimensions, with bony articular surfaces preserved in several specimens. *T. roseae* retains a dermal skeleton, with large unjointed dermal rays, the lepidotrichia and a scaly cover for the humerus. However, like tetrapods, the lepidotrichia are reduced, and the distal endoskeleton is expanded, having several transverse joints at both the level of the proximal carpus and the distal carpus of tetrapods. What is most interesting is the parallel between the loss of the lepidotrichia and the enhanced mobility of the fin joints; the greatest reduction of the fin rays occurs proximally in the regions over the elbow and presumptive wrist joints, not distally as previously supposed. The origin of joints capable of extensive flexion and extension involved not only the origin of facets on the endochondral bones but also the loss of the rays that would have covered them. The phylogenetic position of *T. roseae* reinforced the idea that the reconstruction of the pectoral fin of *Panderichthys rhombolepis* lacking radials was incorrect; indeed, recent analyses have shown that it has a complement of radials that articulate with the ulnare and intermedium[31].

Fossil and genetic data lead to the following picture of the sequence of changes associated with the origin of limbs. The diverse range of fin skeletons seen in fish is generated by a conserved set of genetic tools that are common to all fish that have paired appendages, in particular the early phases of *Hox* gene expression. The adaptive radiation of fish in Devonian aquatic ecosystems, including both near-shore habitats and shallow freshwater streams, was composed of a variety of appendage designs (Fig. 2). One lineage in this radiation acquired fins with an appendage composed of humerus, radius, ulna, ulnare and intermedium and distal bones. One of these lineages, exemplified by *T. roseae*, reduced the dermal rays and expanded the distal joints during the evolution of appendage-based support for the body. Complete loss of the dermal skeleton is associated with a true digital array first seen in Devonian taxa such as *Acanthostega gunnari*. The open question is the extent to which these last steps may be associated with changes to the enhancers that drive phase 2 *Hox* expression in limbs.

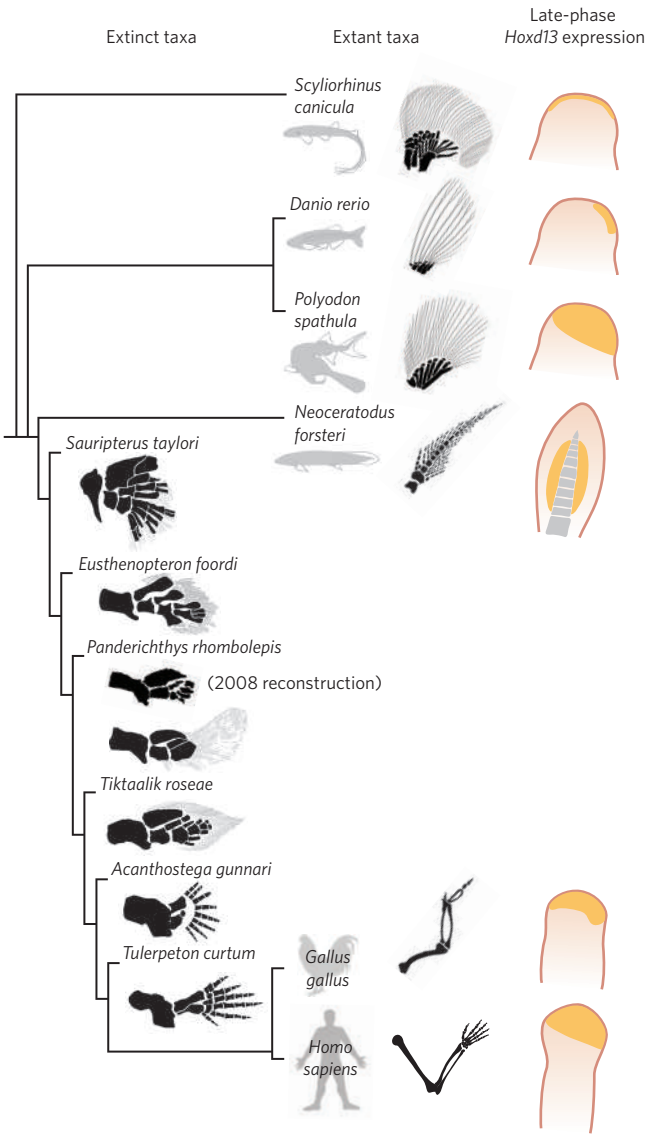


Figure 2 | Deep homology of late-phase *Hox* expression. A phylogenetic tree of gnathostomes, showing the pectoral fins of relevant extant and fossil taxa with known examples of late-phase expression of *Hoxd13* (yellow). Late-phase expression is a primitive feature to sarcopterygians, but its details differ among finned taxa. Basal actinopterygians and lungfish have broad expression domains that resemble the autopodial expression in tetrapods. (Fin illustrations courtesy of K. Monoyios.)

Deep homology and the origin of fins

These advances in understanding the assembly of the tetrapod limb push the issue of limb origins further back in time. How did fins, the foundation for the tetrapod limb, originate? Here, too, deep homology allows predictions to be tested. Both arthropod appendages and tetrapod limbs develop as outgrowths of the body wall that are patterned along three axes — the proximodistal, anteroposterior and dorsoventral axes — often using homologous genes to establish the ordinate axes. If the extraordinary similarities between arthropod and vertebrate appendages reflect the homology of patterning mechanisms, those mechanisms should be present in basal deuterostomes. But limbs are not present in the taxa intermediate between arthropods and vertebrates in deuterostome phylogeny. The clear prediction is that continuity should be found in the outgrowth-patterning mechanisms in arthropods and vertebrates in some other vertebrate structure that predates the origin of fins. Branchial arches and nervous systems are likely candidates, as both are patterned along three major axes, with the former developing as small outgrowths in which many components of limb-patterning circuits are deployed. If this scenario is true, then fins would have arisen by the co-option and ectopic deployment of outgrowth-promoting circuits at novel anatomical sites.

The ideal way to explore this hypothesis would be to examine the state of the limb-development program in finned taxa and in closely related primitively finless taxa. Such comparisons are not possible, however, as the latter are long extinct. But there is now an emerging model within a group of insects that offers the opportunity to reconstruct the more recent origin of a novel outgrowth that seems to have arisen by the co-option of a limb-outgrowth program.

The origin of beetle horns

The scarab beetles (superfamily Scarabaeoidea), which encompass stag beetles (Lucanidae), dung beetles (Scarabaeinae), rhinoceros beetles (Dynastinae) and several other families, include thousands of species with horns of different sizes and shapes that extend from dorsal segments of the head or from the pronotum. In some species, the horns are larger than any of the other appendages, such as the antennae, mouthparts and legs, and can account for up to 30% of the body weight. These extravagant beetles have attracted collectors for centuries, notably Darwin and Alfred Russel Wallace, who each brought new species back from their travels.

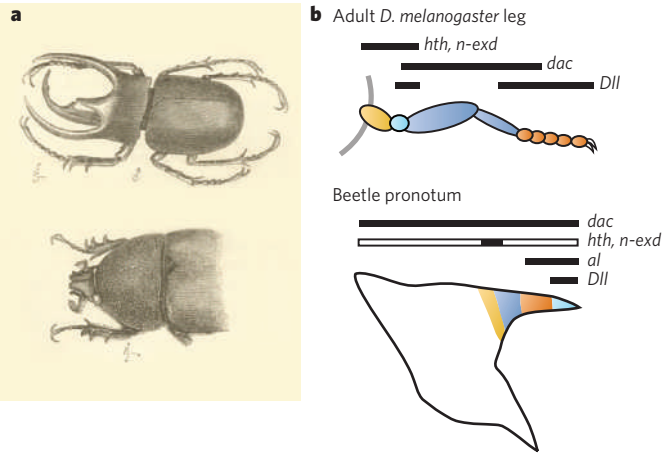


Figure 3 | The evolution of beetle horns by co-option of a limb-outgrowth program. a, Illustration from Darwin's *The Descent of Man* showing the male (top) and female (bottom) forms of the Atlas beetle, *Chalcosoma atlas*. Note the large horns extending from the head. **b**, Similarities in the expression of limb proximodistal axis-patterning gene expression in the adult leg of *Drosophila melanogaster* (top) and the developing horns of *Onthophagus* dung beetles (bottom). Coloured regions denote segments of the leg and regions of the developing horn. Black bars denote gene-expression patterns, indicating the relationship between the expression domains of individual genes in the developing outgrowths. The white bar shows where the listed proteins are co-expressed at only low levels. The expression of outgrowth-promoting genes in beetle horns indicates that these structures evolved by the co-option of an ancient outgrowth program and its deployment at novel anatomical sites. (Panel **a** reproduced from ref. 32; panel **b** modified, with permission, from ref. 35.)

The aptly named Atlas beetle (Fig. 3a) prompted Darwin to remark in *The Descent of Man*[32]: "if we could imagine a male … with its polished bronze coat of mail, and its vast complex horns, magnified to the size of a horse, or even of a dog, it would be one of the most imposing animals in the world."

The Atlas beetle and other extreme species featured prominently in Darwin's conception of sexual selection. Horn formation is typically sexually dimorphic, with males producing large horns and females producing much smaller horns or none at all. Among males of a given species, horn size often varies considerably relative to body size such that small males may not produce horns at all. In Darwin's time, no one had seen horns being used in male combat, so he concluded that they were ornaments used to attract females. But beetle horns are weapons, and there is now extensive evidence that beetles with longer horns prevail in combat with rival males and have a fitness advantage[33].

Co-option of a limb program

Horns are tubular, cuticular projections from the body wall, rather like insect appendages such as antennae, legs and mouthparts. However, unlike typical appendages, horns are unjointed, lack muscles and nerves, and arise from parts of the body where insects do not generally develop outgrowths, so they lack obvious homology to appendages. Most beetle species are hornless, including the vast majority of scarab species; therefore, horns are clearly evolutionary novelties. So how did they arise? The key to answering this question is understanding the developmental basis of horn formation.

Recent studies of horn development in the dung beetle *Onthophagus* have shown that they form from compact discs of epidermal cells that proliferate during the late larval period and then evert to their full length during the pupal moult, just as typical body appendages do[30]. Most strikingly, developing horns express a suite of genes that subdivide the proximodistal axis of typical insect limbs, including expression of the *Distal-less* (*Dll*) gene in the distal tip of the developing horn and expression of the *homothorax* (*hth*) gene and the nuclear form of the Extradenticle protein (encoded by *exd*) in the most proximal base of the developing horn[34–36] (Fig. 3b). These results suggest that beetle horns are the product of the co-option and deployment of an outgrowth program at novel anatomical sites.

An inordinate fondness for beetle horns?

The widespread distribution and great diversity of horns has prompted many naturalists to consider the evolution of horns in scarabs[37]. This raises the question of whether horns have arisen independently numerous times and, if so, whether this has occurred by the same mechanism. Although analyses of the distribution of horns in adults have suggested numerous independent gains, even within a single genus, recent studies of development in larvae have shown that the potential to make horns is widespread, even among hornless species. Armin Moczek and colleagues have shown that although certain members of the dung beetle genus *Onthophagus* develop pronotal horns during the late larval period and retain them to adulthood[38], many other species transiently develop horns as larvae and then resorb them in the pupal period. Furthermore, these authors uncovered evidence that the larval horns have an important role in the splitting of the head capsule during the larval-to-pupal moult.

These findings explain why the capacity to make adult horns has been widely maintained and could therefore account for the multiple independent gains (and losses) of adult horns in this genus. Furthermore, because the adult pronotal horns are clearly derived from larval horns, the findings suggest that adult *Onthophagus* horns did not evolve from scratch for male combat but are exaptations (structures that first originated to serve another, unrelated, function).

Horn development has been studied less in other clades, so it is not clear how often horns have arisen. However, it has been suggested that their presence at distinct anatomical sites on the head and the pronotum, which are each associated with deployment of proximodistal axis-patterning genes[36], results from at least two independent co-options of an outgrowth-promoting program[39].

Deep homology and parallel evolution

Darwin closed *The Descent of Man*, his second greatest work, with the line: "Man still bears in his bodily frame the indelible stamp of his lowly origin." These words have never rung more true. Researchers are now starting to appreciate the presence of far more indelible stamps of humanity's lowly metazoan origins than Darwin could ever have imagined. The detection of deep homologies offers more than new glimpses of evolutionary history, however. Such homologies provide a profound insight into the evolutionary process. Studies of deep homology are showing that new structures need not arise from scratch, genetically speaking, but can evolve by deploying regulatory circuits that were first established in early animals. But herein lies a challenge for the next generation of biologists: if the mechanisms behind the formation of diverse organs are ancient and highly conserved, then parallel evolution must be considered a fact of life in the phylogenetic history of animals[2].

With the growth of developmental genetics, it is possible to see beyond the view of homologies working at the level of whole organs. The mechanisms that define the ordinate axes of structures, the genetic circuits that pattern them, and the cell types with which organs are formed can be considered. The more that researchers look, the more they will find that the same tools have been used to build a great variety of structures long thought to have independent histories[40–42]. Discerning what has been conserved and what is novel in the origins of organs and body plans will be possible only with more comparative data, experiments on non-model animals, and targeted fossil discoveries from crucial nodes in the tree of life. ■

1. Shubin, N., Tabin, C. & Carroll, S. Fossils, genes, and the evolution of animal limbs. *Nature* **388,** 639–648 (1997).
This paper provides an original description of deep homology and an analysis of the extensive developmental similarities between arthropod and vertebrate appendages.
2. Gould, S. J. *The Structure of Evolutionary Theory* (Harvard Univ. Press, 2002).
Gould's magnum opus has an extensive analysis of deep homology and the fundamental ways that parallel evolution is a major pattern of evolution.
3. Panganiban, G. *et al.* The origin and evolution of animal appendages. *Proc. Natl Acad. Sci. USA* **94,** 5162–5166 (1997).

4. Mercader, N. *et al.* Conserved regulation of proximodistal limb axis development by Meis1/Hth. *Nature* **402,** 425–429 (1999).
5. Capdevila, J. *et al.* Control of vertebrate limb outgrowth by the proximal factor Meis2 and distal antagonism of BMPs by Gremlin. *Mol. Cell* **4,** 839–849 (1999).
6. Darwin, C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (John Murray, 1859).
7. Salvini-Plawen, L. V. & Mayr, E. On the evolution of photoreceptors and eyes. *Evol. Biol.* **10,** 207–263 (1977).
8. Halder, G., Callaerts, P. & Gehring, W. J. Induction of ectopic eyes by targeted expression of the eyeless gene in *Drosophila*. *Science* **267,** 1788–1792 (1995).
9. Gehring, W. J. New perspectives on eye development and the evolution of eyes and photoreceptors. *J. Hered.* **96,** 171–184 (2005).
10. Kozmik, Z. *Pax* genes in eye development and evolution. *Curr. Opin. Genet. Dev.* **15,** 430–438 (2005).
11. Oakley, T. H. The eye as a replicating and diverging, modular developmental unit. *Trends Ecol. Evol.* **18,** 623–627 (2003).
12. Arendt, D., Tessmar-Raible, K., Snyman, H., Dorresteijn, A. W. & Wittbrodt, J. Ciliary photoreceptors with a vertebrate-type opsin in an invertebrate brain. *Science* **306,** 869–871 (2004).
By showing that polychaetes have a 'vertebrate-type' opsin in their brains, this paper reveals how divergent eyes of bilaterians could have evolved from a common ancestor.
13. Erclik, T., Hartenstein, V., Lipshitz, H. D. & McInnes, R. R. Conserved role of the *Vsx* genes supports a monophyletic origin for bilaterian visual systems. *Curr. Biol.* **18,** 1278–1287 (2008).
14. Pearse, J. S. & Pearse, V. B. Vision of cubomedusan jellyfishes. *Science* **199,** 458 (1978).
15. Yamasu, T. & Yoshida, M. Fine structure of complex ocelli of a cubomedusan, *Tamoya bursaria*. *Cell Tissue Res.* **170,** 325–339 (1976).
16. Kozmik, Z. *et al.* Assembly of the cnidarian camera-type eye from vertebrate-like components. *Proc. Natl Acad. Sci. USA* **105,** 8989–8993 (2008).
17. Suga, H., Schmid, V. & Gehring, W. J. Evolution and functional diversity of jellyfish opsins. *Curr. Biol.* **18,** 51–55 (2008).
18. Plachetzki, D. C., Degnan, B. M. & Oakley, T. H. The origins of novel protein interactions during animal opsin evolution. *PLoS ONE* **10,** e1054 (2007).
19. Coates, M. I., Jeffery, J. E. & Ruta, M. Fins to limbs: what the fossils say. *Evol. Dev.* **4,** 390–401 (2002).
20. Shubin, N. The evolution of paired fins and the origin of the tetrapod limb: phylogenetic and transformational approaches. *Evol. Biol.* **28,** 39–86 (1995).
21. Shubin, H. Daeschler, E. B. & Jenkins, F. A. The pectoral fin of *Tiktaalik roseae* and the origin of the tetrapod limb. *Nature* **440,** 764–771 (2006).
The description of the pectoral fin and shoulder of *T. roseae* shows how the closest relatives to tetrapods evolved supporting appendages using components of the tetrapod autopod.
22. Sordino, P. & Duboule, D. A molecular approach to the evolution of vertebrate paired appendages. *Trends Ecol. Evol.* **11,** 114–119 (1996).
23. Sordino, P., Hoeven, F. V. D. & Duboule, D. *Hox* gene expression in teleost fins and the origin of vertebrate digits. *Nature* **375,** 678–681 (1995).
24. Zakany, J., Fromental-Ramain, C., Warot, X. & Duboule, D. Regulation of number and size of digit by posterior *Hox* genes: a dose dependent mechanism with potential evolutionary implications. *Proc. Natl Acad. Sci. USA* **94,** 13695–13700 (1997).
25. Wagner, G. P. & Chiu, C.-H. The tetrapod limb: a hypothesis on its origin. *J. Exp. Zool.* **291,** 226–240 (2001).
26. Spitz, F., Gonzalez, F. & Duboule, D. A global control region defines a chromosomal regulatory landscape containing the *HoxD* cluster. *Cell* **113,** 405–417 (2003).
27. Davis, M. C., Dahn, R. D. & Shubin, N. H. An autopodial-like pattern of *Hox* expression in the fins of a basal actinopterygian fish. *Nature* **447,** 473–476 (2007).
28. Johanson, Z. *et al.* Fish fingers: digit homologues in sarcopterygian fish fins. *J. Exp. Zool.* **308,** 757–768 (2007).
29. Ahn, D.-G. & Ho, R. K. *Hox* genes and development of paired fins in teleost: an alternative view. *Dev. Biol.* **295,** 419–435 (2008).
30. Freitas, R., Zhang, G. & Cohn, M. Biphasic *Hoxd* gene expression in shark paired fins reveals an ancient origin of the distal limb domain. *PLoS ONE* **2,** e754 (2007).
31. Boisvert, C. A., Mark-Kurik, E. & Ahlberg, P. The pectoral fin of *Panderichthys* and the origin of digits. *Nature* **456,** 636–638 (2008).
32. Darwin, C. *The Descent of Man, and Selection in Relation to Sex* (John Murray, 1871).
33. Hunt, J. & Simmons, L. W. Status-dependent selection in the horn-dimorphic beetle *Onthophagus taurus*. *Proc. R. Soc. Lond. B* **268,** 2409–2414 (2001).
34. Moczek, A. P. & Nagy, L. M. Diverse development mechanisms contribute to different levels of diversity in horned beetles. *Evol. Dev.* **7,** 175–185 (2005).
35. Moczek, A. P. Integrating micro- and macroevolution of development through the study of horned beetles. *Heredity* **97,** 168–178 (2006).
36. Moczek, A. P., Rose, D., Sewell, W. & Kesselring, H. R. Conservation, innovation, and the evolution of horned beetle diversity. *Dev. Genes Evol.* **216,** 655–665 (2006).
37. Arrow, G. H. *Horned Beetles* (Junk, 1951).
38. Moczek, A. P., Cruickshank, T. E. & Shelby, A. When ontogeny reveals what phylogeny hides: gain and loss of horns during development and evolution of horned beetles. *Evolution* **60,** 2329–2341 (2006).
39. Emlen, D. J., Lavine, L. C. & Ewen-Campen, B. On the origin and evolutionary diversification of beetle horns. *Proc. Natl Acad. Sci. USA* **104,** 8661–8668 (2007).
40. Denes, A. S. *et al.* Molecular architecture of annelid nerve cord supports common origin of nervous system centralization in bilateria. *Cell* **129,** 277–288 (2007).
41. Pueyo, J. I., Lanfear, R. & Couso, J. P. Ancestral Notch-mediated segmentation revealed in the cockroach *Periplaneta americana*. *Proc. Natl Acad. Sci. USA* **105,** 16614–16619 (2008).
42. Tessmar-Raible, K. *et al.* Conserved sensory-neurosecretory cell types in annelid and fish forebrain: insights into hypothalamus evolution. *Cell* **129,** 1389–1400 (2007).

Acknowledgements N.S. is supported by grants from the National Science Foundation and the National Geographic Society. C.T. is supported by grants from the National Institutes of Health. S.C. is an investigator of the Howard Hughes Medical Institute.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence should be addressed to N.S. (nshubin@uchicago.edu).

NATURE|Vol 457|12 February 2009|doi:10.1038/nature07892

The *Beagle* in a bottle

Angus Buckling[1], R. Craig Maclean[1], Michael A. Brockhurst[2] & Nick Colegrave[3]

Why infer evolution when you can watch it happen in real time? This is the basic premise of using populations of fast-replicating microorganisms in test tubes to study evolution. The approach, known as experimental evolution, has provided a way of testing many of the key hypotheses that arose from the modern evolutionary synthesis. However, details of the unnatural histories of microorganisms in test tubes can be extrapolated only so far. Potential future directions for the approach include studying microbial evolution for its own sake under the most natural conditions possible in the test tube, and testing some qualitative theories of genome evolution.

Charles Darwin's insights into the nature and mechanism of the evolutionary process were built on careful observation of the patterns of variation across the natural world. In *On the Origin of Species*[1], Darwin brought together evidence from an incredible diversity of taxa to support his case. But one major group notably lacking from this work, the microorganisms, has since provided some of the best experimental support for the theory of evolution by natural selection[2]. In this Review, we examine how microorganisms, and in particular evolution experiments using microorganisms, have helped to unlock the details of the adaptive process. In particular, we highlight some of the most important insights that this approach has provided, those that we think would have excited Darwin, and we discuss the important role that this approach will continue to have in furthering the understanding of organic diversity.

When Darwin developed his theory of natural selection, he assumed that evolutionary change would usually be slow and, in general, not directly observable. As he wrote in *On the Origin of Species*: "We see nothing of these slow changes in progress, until the hand of time has marked the long lapses of ages, and then so imperfect is our view into the long past geological ages that we only see that the forms of life are now different from what they were."

The fossil record was valuable in supporting his thesis, but it was also imperfect and open to multiple interpretations. However, one of Darwin's crucial realizations was that the process of natural selection had an analogue in the artificial selection applied by breeders of animals and plants. Darwin frequently used examples from fields such as pigeon breeding to illustrate the power of cumulative selection to generate rapid change in biological organisms. The value of such indirect evidence into the power of selection is indisputable, but how much more powerful is it to observe natural selection directly and in real time?

What is experimental evolution?

Experimental evolution has transformed evolutionary biology from a historical science, in which unseen processes are inferred from evolutionary end points, to one in which evolution can be studied in real time. The method involves propagating replicate populations of living things in defined laboratory environments for tens, hundreds or even tens of thousands of generations. The experimenter controls the environmental conditions but does not directly impose the selection, as is the case with pigeon breeding. Instead, selection results from the "struggle for existence" between individuals within each population; in other words, selection is natural. A wide range of organisms have been studied in this way, including nematodes, fruitflies and mice[2], but by far the easiest and most productive organisms to work with are microorganisms, primarily bacteria, yeast and viruses. Their large populations and short generation times favour rapid evolution, and the ability to store populations cryogenically in suspended animation allows direct comparisons to be made between evolved, ancestral and intermediate forms, providing researchers with a 'living fossil record'[3]. Moreover, many microorganisms have relatively simple and well understood genomes, allowing both genetic manipulation and, to some extent, identification of the genetic targets for selection.

Perhaps the first experimental evolutionist was William Dallinger, a contemporary of Darwin. Using equipment not dissimilar to that seen in modern-day microbiology laboratories, Dallinger maintained populations of flagellates under conditions of gradually increasing temperature[4]. After several thousand generations, his flagellates had evolved to grow at temperatures far exceeding the thermal range of their ancestors, providing a dramatic demonstration of the power of natural selection to adapt organisms to changing environmental conditions. It is an interesting historical note that Darwin, although fully aware and supportive of this work, does not seem to have appreciated its value as evidence in support of his theory.

Despite this long history, the field of microbial experimental evolution did not take off until the early 1990s, following the publication of work by Richard Lenski and colleagues[3] on the long-term adaptation of the bacterium *Escherichia coli* to a novel laboratory environment. This experiment contains many elements common to experimental evolution studies and is outlined in Box 1. This work led to a rapid rise in the popularity of microbial experimental evolution as a method for studying evolution (Fig. 1), and it has been used to address questions spanning the entire spectrum of the modern evolutionary synthesis, from the tempo and mode of adaptation and diversification to the evolution of sex and sociality.

What has been learned

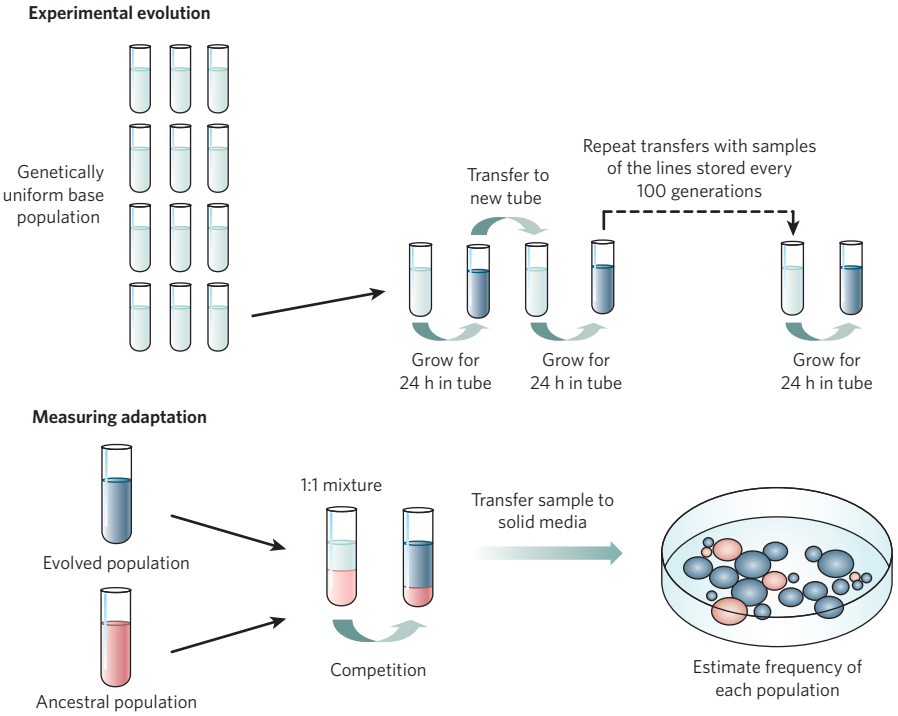
Experimental evolution has provided a wealth of information about the way evolution works by allowing researchers to test their theories directly. Here we consider some of the key findings that sprang from this work.

The importance of chance in the adaptive process

Darwin proposed that natural selection would lead to adaptation and organisms that are well fitted to their environment. The incorporation of genetics into Darwin's evolutionary framework (the 'modern evolutionary synthesis')[5] led to a number of explicit predictions concerning the dynamics of adaptation, many of which have now been confirmed by

[1]Department of Zoology, University of Oxford, Oxford OX1 3PS, UK. [2]School of Biological Sciences, University of Liverpool, Liverpool L69 7ZB, UK. [3]School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, UK.

Box 1 | The long-term selection experiment



On 15 February 1988, Richard Lenski started a selection experiment that continues to this day. The experimental design is startlingly simple: 12 populations of the bacterium *Escherichia coli* B were grown in identical flasks containing a simple liquid medium with glucose as the sole carbon source. Every day, 1% of each population was transferred to a fresh flask (see box figure, top). For *E. coli* B, this was a new environment, and one to which it was initially poorly adapted. Each 24-hour growth cycle allowed 6.6 bacterial generations, and after every 100 generations a sample of each population was stored in the freezer alongside the ancestral *E. coli* B strain. Because the 12 populations were initially genetically identical, the only source of variation for natural selection to act on was *de novo* mutation.

The adaptation of each population to the environment was assessed by directly comparing the evolved bacteria to the *E. coli* B ancestor in head-to-head competition experiments under exactly the same conditions as one 24-hour growth cycle of the selection experiment (see box figure, bottom). Samples of the evolved and ancestral populations (the latter with a genetic colour marker) were removed from the freezer and grown individually. Small numbers of bacteria from both populations were then mixed in a 1:1 ratio in the same tube and allowed to compete for 24 hours. Samples of the mixture were then grown on solid media that allow colonies derived from evolved and ancestral cells to be distinguished on the basis of their colour. The change in the ratio of the types was then used to estimate the fitness of the evolved lines relative to the ancestor.

In 1994, the results of 10,000 generations of the selection experiment were published[7]. This revealed that the relative fitness of the 12 experimental populations had increased by approximately 40% but that most of this increase occurred in the first 2,000 generations of evolution. Cell size almost doubled during 10,000 generations, following a similar trajectory to relative fitness, with the most rapid increase over the first 2,000 generations. Importantly, the evolutionary trajectories of both relative fitness and cell size, and the relationship between them, were subtly different between the replicate populations, suggesting that chance and contingency, as well as natural selection, had important roles in the evolution of these populations.

experimental evolution. One particularly robust example is the way the rate of adaptation changes through time: fitness initially increases rapidly as mutations that have a large effect become fixed, and then plateaus as beneficial mutations become rarer and have a smaller effect[6]. One of the reasons for this consistency between theory and experiments is the large population sizes of microorganisms studied, which increases the predictability of the evolutionary process. However, chance does, surprisingly, still have an important effect, even in massive populations. For example, initially identical populations of *E. coli* placed in identical, simple environments follow different evolutionary trajectories, in terms of both fitness itself and, even more strikingly, in phenotypic traits such as cell size[7] (Box 1). Genetic analyses of these populations have shown that, by chance, different random mutations, albeit often in the same gene, have been fixed in different populations, subtly altering the trajectory of evolution[8,9]. The fact that subsequent mutations were then contingent on prior mutations (epistasis) acted to reinforce this divergence[10].

The ability to evolve may itself evolve

Natural selection operates on genetic diversity, so it follows that mechanisms that generate diversity may themselves be under selection. Indeed, Darwin recognized that sexual reproduction may be an example of an evolvability mechanism[1], as it allows more rapid adaptation of populations. Because the benefits of sex can take considerable time to be realized in multicellular organisms, facultatively sexual microorganisms (such as yeast[11] and single-celled algae of the genus *Chlamydomonas*[12]), provide a unique opportunity to determine the conditions under which sex is advantageous. Whether the organisms have sex can be manipulated by the experimenter in a variety of ways, and otherwise identical sexual and asexual populations are allowed to evolve under defined conditions. The main conclusion from these studies is that sex does indeed favour more rapid adaptation but only when populations are both large and poorly adapted to their environment[11,12]. In most cases, the key benefit of sex is that it brings together beneficial mutations in the same individual that would otherwise be competing against each other in different individuals, an idea originally proposed in the 1930s[13,14].

A more fundamental mechanism of evolvability is mutation itself. In a well-adapted organism, most mutations are likely to be deleterious, so it might be expected that natural selection drives mutation rates down until the physiological costs of error correction become too great[15]. However, for asexual microorganisms, increased mutation rates, which generally result from dysfunctional DNA proof-reading enzymes[16], may confer an indirect advantage by generating a larger number of beneficial mutations that

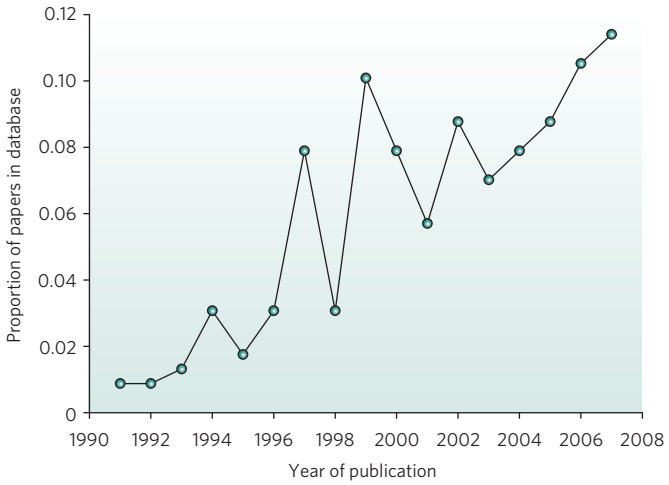


Figure 1 | The rise of experimental evolution. We constructed a bibliographical database of experimental evolution papers by searching the ISI Web of Knowledge for papers that contain the keywords 'experimental evolution' or 'selection experiment' and were published between 1991 and 2007. This search was then manually edited to produce a database of experimental evolution papers involving microorganisms. The reduced database contains 268 papers, which is an underestimate of the total size of the literature in the field, but we have no compelling reason to think that this database is not a random sample of the experimental evolution literature. The graph shows the proportion of the papers in this reduced database that were published in each year.

effectively offset the costs of the deleterious ones[17]. Crucially, bacteria with increased mutation rates ('mutators') have been observed to evolve *de novo* in the laboratory[18–20]. As with sexual reproduction, increased mutation rates are typically favoured in populations that are poorly adapted to their environment. However, unlike sex, increased mutation rates are more beneficial in small populations that would otherwise spend longer than large populations waiting for the next beneficial mutation[21].

Diversity is the rule and not the exception

It was the great diversity of life on Earth that inspired Darwin to develop his theory of natural selection, and the question of why there are so many species continues to fascinate evolutionists to this day. In an early example of experimental ecology, Georgii Gause's studies[22] on coexistence of competing species of the unicellular ciliate *Paramecium* gave rise to the principle of competitive exclusion, which states that diversity cannot be maintained without environmental complexity in the form of multiple ecological niches. Experimental evolution has extended this finding over evolutionary timescales, showing that the amount of diversity that evolves from initially genetically uniform populations increases as a function of environmental complexity, in terms of spatial heterogeneity[23] (Fig. 2) or types of food resource[24]. Furthermore, these studies show that fitness trade-offs between different ecological niches readily evolve and are crucial if diversity is to be stably maintained. However, what is perhaps more surprising is that stable diversity, and hence specialization, can also evolve in apparently homogeneous environments containing a single food resource[25–29]. This diversity results from trade-offs between primary resource and waste metabolite utilization[27], but more fundamental biochemical trade-offs between the rate and yield of energy production are also likely to be important[30]. The main conclusion of these studies is that environmental complexity influences diversity, but given enough time and competition for resources, diversity is always likely to evolve.

Costly social action can be explained by kin selection

Like all organisms, microorganisms have social lives, with individual cells carrying out a range of extracellular actions that can affect the fitness of nearby cells[31]. These include the formation of multicellular structures and the production of enzymes and chelators to release nutrients from the environment[31,32]. However, many of these social actions are metabolically costly to individual cells yet provide a benefit to the group, and as such are open to exploitation by social 'cheats' who reap the rewards of social action without paying the costs. In *The Descent of Man*[33], Darwin suggested that individually costly social action could evolve if the benefits accrued to family members, an evolutionary process later formalized by W. D. Hamilton as kin selection[34]. By manipulating the genetic relatedness of groups of interacting cells, evolution experiments using a wide range of microorganisms and their social traits have shown that cooperation only persists when genetic relatedness is high[31]. For example, the opportunistic pathogen *Pseudomonas aeruginosa*, like most bacteria, produces extracellular iron-scavenging molecules, known as siderophores. Social cheats that produce fewer siderophores but can still use the iron-bound siderophores produced by others readily evolve and increase in frequency from wild-type populations[35] (Fig. 3). High siderophore production can be maintained only when populations are exposed to strong genetic bottlenecks, ensuring that relatedness is high[36]. Moreover, this study[36] also confirmed theoretical predictions that competition between groups (allowing productive groups to produce more propagules) is crucial for the evolution of costly cooperation. There are numerous other examples of the role of kin selection in explaining cooperation in microorganisms, most notably that of social amoebae (such as slime moulds of the genus *Dictyostelium*) sacrificing themselves to form stalks that support viable spores[37].

Species co-evolve

Darwin recognized the importance of co-evolution, which is the reciprocal evolution of species or populations in response to each other. He saw it in particular as an explanation for the good fit of the length of pollinators' tongues to the nectar spurs of the flowers they visit[1]. Co-evolution is likely to be a powerful force because a change in one species alters selection on the other species at a much faster rate than that typically caused by changes in the physical environment[38]. The most extreme example of rapidly changing selective pressures is that resulting from the interaction between hosts and their lethal parasites. The evolution of parasite-resistant hosts from a previously sensitive population will drastically change the parasite's lot from utopian to the brink of extinction, and likewise for the host population if the parasites evolve to overcome this resistance. Because of the potential speed of host–parasite co-evolution, it has been implicated as a key force in a range of evolutionary processes, including speciation[38] and the evolution of sex[39]. Yet it has proven to be one of the most elusive of evolutionary phenomena to detect in natural populations. Studies using microorganisms, primarily bacterial hosts and parasitic viruses (Fig. 4), have allowed co-evolution to be directly observed by taking advantage of the living fossil record[40–43]. Populations can be assayed against past, contemporary and even future enemy populations, allowing the reciprocal evolution of resistance and infectivity to be directly measured. These studies not only have confirmed the rapidity and persistence of co-evolutionary change but also have allowed tests of the consequences of co-evolution. For example, co-evolution can radically alter host genetic diversity[44] and impose selection for evolvability (increased mutation rates)[20]. Experimental studies that focus on the evolution of mutualisms (specifically, in *E. coli* and plasmids that carry antibiotic resistance genes) have shown that such mutualistic interactions become more and more specialized over evolutionary time[45], confirming Darwin's hunch.

Criticisms and caveats

Despite the success of microbial experimental evolution in providing crucial tests of evolutionary theory, there are inevitably limitations. In this section, we examine five of the most common criticisms.

Test tubes are unnatural environments

It may be difficult to show how evolution works in nature by studying it in an unnatural environment, such as a test tube. However, although a test tube is unnatural at the beginning of the study, after the hundreds or thousands of generations in a typical study, it becomes a natural environment with respect to the microorganism's recent evolutionary history[3]. By comparison, many of the organisms so beloved of researchers investigating

evolution in the wild are relative newcomers to their current environments, as is the case for the well-studied Soay sheep population that was introduced to the Scottish island of St Kilda about 4,000 years ago with human settlers[46]. Thus, a great advantage of evolution in test tubes is that the organisms have evolved in precisely the same environment in which they are subsequently assayed. Compared with taking organisms out of their environment and measuring them under laboratory conditions, the advantages of this approach are clear.

Test tubes are too simple to examine real-world complexity

The criticism that a test tube cannot mirror the complexity of natural environments misses a key rationale behind this approach. Experimental evolution does not seek to mimic specific systems in all their complexity. Instead, the microcosms are biological models in which researchers attempt to capture the essence of evolving systems in general and so can shed light on general processes that are expected to occur in all life on Earth. Evolutionary biologists have successfully used simple mathematical models to provide insight into fundamental evolutionary phenomena. No one would argue, for example, that because R. A. Fisher's fundamental theorem[13] does not contain parameters for every aspect of the biology of even a single species, it can suggest nothing useful about evolution. Simplicity is a strength, rather than a weakness. Experimental evolution can provide support for a given theory, but it indicates little about its relative importance in explaining patterns in nature where many selective forces are likely to act simultaneously. Experiments can be made more complex, with many interacting selective forces included[47], to provide a clearer picture of the ecological and genetic mechanisms studied. Paradoxically, however, such complexity may not be the most productive use of experimental evolution because, by definition, generality is lost. Again, the analogy with theoretical models is clear: the more parameters a model has, the less generally useful it becomes.

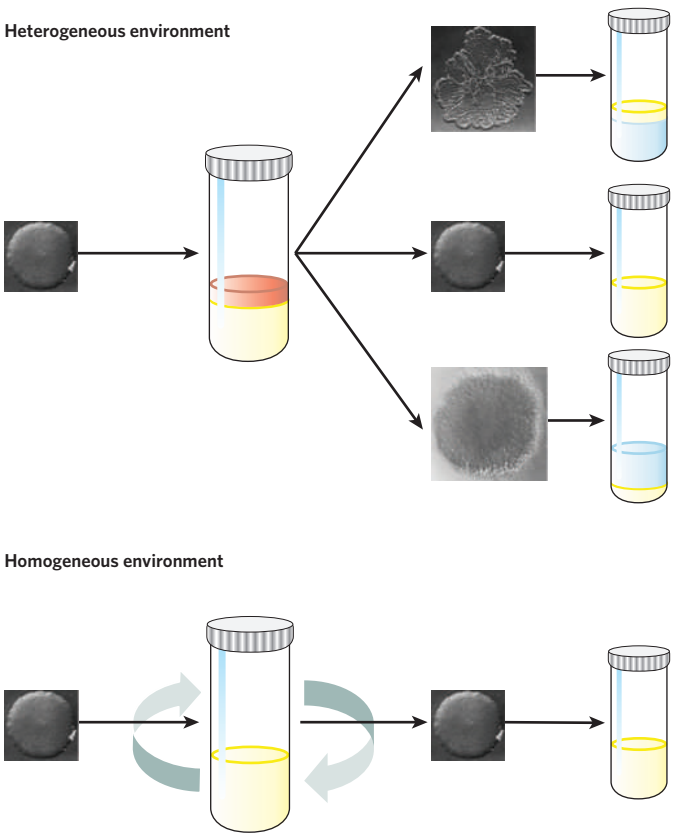


Figure 2 | Adaptive radiation in a heterogeneous environment. Paul Rainey and Mike Travisano placed a single clone of the bacterium *Pseudomonas fluorescens* into two different environments, one with and one without spatial heterogeneity[23]. After several days, the bacteria grown in each environment were plated onto solid medium. The colony grown in the presence of spatial heterogeneity gave rise to a diverse range of colony morphologies, whereas the colony grown in its absence did not. Crucially, these different colony morphotypes were each adapted to occupy a different region of the liquid broth (right; yellow band). This finding was tested by re-inoculating each morphotype into a fresh sterile microcosm of liquid broth. (Images reproduced, with permission, from ref. 23.)

Microorganisms are different from larger organisms

Certain questions are clearly beyond the reach of experimental evolution using microorganisms, such as the evolution of consciousness. However, biologists unfamiliar with microorganisms may be surprised at how similar the lives of microbiota and macrobiota are. As discussed earlier, some microorganisms have sex (although, to complicate matters, bacteria and viruses transfer genes between species[48]), and all microorganisms have complex social lives, in which they communicate, cooperate and are spiteful to one another[31]. It has recently been shown that they even grow old and decrepit: the poles of bacteria are subject to ageing[49,50], allowing experimental tests of the hypothesized trade-off between reproduction and ageing[51]. The key is to choose an organism that fits the question, as is the case in all areas of evolutionary ecology.

Microevolution cannot be scaled up

Several experimental evolution studies have framed their results in the context of species-level, or macroevolutionary, change. For example, the replacement of one mutation by another, followed by a period of evolutionary stasis, has been interpreted as supporting a punctuated-equilibria model of macroevolution[52], and the rapid diversification of a single bacterial clone into genetically distinct niche specialists has been described as adaptive radiation[23]. Changes in a test tube are largely the result of a few point mutations or insertions[8,53], but they have large phenotypic effects as a result of the novelty of the laboratory environment. This is clearly not macroevolution, because genotypes are not species. But similar processes are likely to be driving microevolutionary and macroevolutionary change. For example, adaptive radiations are almost certainly driven by competition and ecological opportunity[54]; exactly the same mechanisms have been identified in evolution experiments. Furthermore, recent experimental evolution studies have detected the signatures of speciation. For example, the theory of ecological speciation suggests that reproductive isolation evolves as a by-product of divergent ecological selection[54], and this has recently been shown in experimental populations of yeast[55]. Perhaps macroevolution could occur in test tubes after all.

Experimental evolution studies are contrived

The control afforded by experimental evolution can lead to rather inevitable outcomes. This is especially so when experiments are carried out over short timescales using genotypes whose phenotypic characteristics have already been defined in the selective environment. For example, several studies have begun with mixtures of two genotypes, a wild type and an isogenic mutant with a particular gene knocked out, and followed the frequencies of these genotypes for a few tens of generations. Such studies can provide a useful biological demonstration of theory and may be the only feasible way to quantify weak selective forces, but the true strength of these studies is that experiments can be carried out over timescales in which the evolutionary outcomes are not limited by the genetic variation initially present. Such studies allow researchers to examine processes that occur beyond the immediate short term and that are not easily examined in other ways.

The future

We have discussed how experimental evolution studies have provided insights into phenotypic evolution in real time and the genetic changes underlying it. But what might such studies show about the evolution of genomes themselves? And can they be applied to solve problems in the real world?

Experimental evolution at a genomic scale

The specimens catalogued so carefully by Victorian naturalists, which inspired Darwin's ideas, have been replaced by a new kind of specimen:

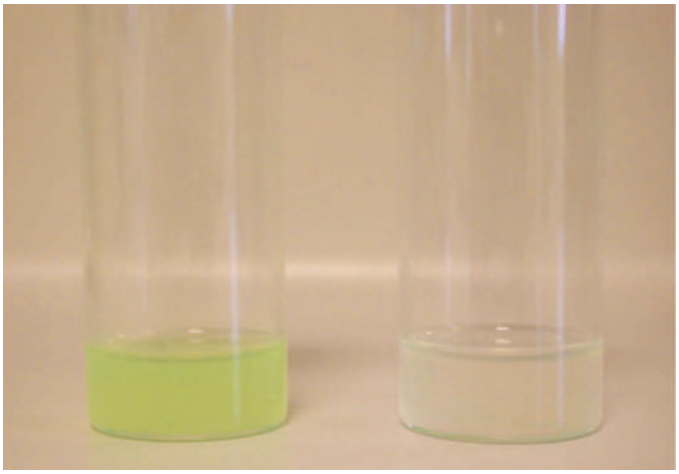


Figure 3 | Cooperation and cheating in *Pseudomonas aeruginosa*. The tube on the left shows a wild-type cooperating population of the bacterium *P. aeruginosa* producing pyoverdine, an extracellular, iron-scavenging molecule (or siderophore) that is a yellow-green colour. The tube on the right shows a pyoverdine-negative 'cheat', which can exploit the siderophore but does not produce it.

the genome sequence. The ability to sequence genomes, and measure transcriptomes and proteomes, increases almost daily, and computational developments in metabolic-network analysis are increasingly resulting in a change away from simple observation towards hypothesis-driven questions[56–58]. What determines the physical order of genes in the genome? When do genomes evolve to be robust to mutation? How do networks of genetic regulation evolve? What is the minimal genome for a particular set of environmental conditions? Fortunately, the microorganisms used in experimental evolution studies are cheap and easy to sequence, and it is increasingly likely that it will be possible to carry out large-scale sequencing of not only the end points of selection experiments but also of many individuals in between. Experimental evolution has, in the past, allowed evolution to be observed directly at the phenotypic level; in the future, it will provide a means of examining directly the processes that mould the evolution of genomes themselves. Some of the earlier studies of experimental evolution set out to address these types of question[59], but the technological advances that accompanied the genomic revolution will allow researchers to address them in a much more systematic way.

Despite the incredible potential of experimental evolution as a tool to study the evolution of genomes, we can see significant challenges to the development of this field. Experimental evolution studies succeed when they explicitly test clearly defined and general hypotheses that should apply to any biological system, and the success of 'omic'-scale studies in experimental evolution will depend on their ability to stay within this hypothesis-driven framework. Many experimental evolution studies that investigated the genetic basis of adaptation had the goal of simply identifying the mutations underlying adaptation to a particular laboratory environment. However, the arbitrary and artificial nature of selective environments imposed by researchers makes it questionable how relevant this approach is for understanding evolution. Another challenge will be to understand the functional consequences of genetic variation: producing long lists of mutations that are responsible for adaptation could amount to little more than 'stamp collecting' unless the consequences of mutations are adequately characterized.

Finally, many of the most spectacular examples of adaptation in selection experiments result from modest changes at the genomic level[60–62], suggesting that experiments designed to investigate broad features of genome evolution will need to find ways to accelerate adaptation or find new model systems to work with. Experiments involving thousands of generations might not be sufficient to understand the underlying causes of some of the differences in genome structure and content that are clear from comparing the genomes of species that have diverged from each other over millions of years.

Studying microorganisms for their own sake

Experimental evolution studies using microorganisms are dominated by tests of universal theories, but the approach can also improve knowledge of the evolution and ecology of the microorganisms themselves. This is clearly a useful aim, given the ubiquity of microorganisms and their important role in all ecosystems. However, it is crucial to use nature as a guide. It is most useful to identify phenotypes that are clearly under selection in natural populations, as well as the key selective forces that operate on them. For example, public-goods cheats in the bacterium *P. aeruginosa*[63,64], which produce less siderophore than normal and are often associated with reduced virulence[65,66], are much more prevalent in chronic infections than in the initial colonizing populations[63,64]. Compared with success in the external environment, success within hosts imposes considerably different selective pressures (such as immunity, predation, antibiotics and scale of competition), and experimental evolution studies can determine the relative importance of these forces in creating natural patterns of social behaviour. Using experimental evolution to determine how microorganisms respond to selective pressures in their natural environments will also allow their evolution to be anticipated. This is likely to be particularly useful for testing the propensity to evolve resistance to antibiotics[67].

The next logical step in studying the evolution of microorganisms is to carry out experimental evolution in natural environments. At first this may seem to disregard one of the key advantages of the whole approach: control over the environment. However, 'natural microcosms', such as tubes of soil or pond water, or even isogenic hosts[18], can be replicated, and environmental and genetic variables manipulated. The crucial difference between such studies and traditional experimental evolution is that the importance of the manipulated selective pressure can be determined among all the other natural selective pressures.

Evolving helpful microorganisms

Humans have used selection in a range of organisms, including microorganisms. For example, the lactic-acid-producing bacterium *Streptococcus thermophilus*, which is used in the production of yoghurt, has been artificially selected for thousands of generations, possibly from a human oral commensal ancestor[68], and brewing and bread yeasts have been artificially selected from naturally occurring grape yeasts. Despite this illustrious history of applied selection experiments on microorganisms, selection is not being fully exploited to produce microorganisms that will be useful in medical, environmental or agricultural contexts. For example, bioreactors, which rely on microorganisms to break down or produce chemicals, sometimes fail[69] because of colonization by more competitive

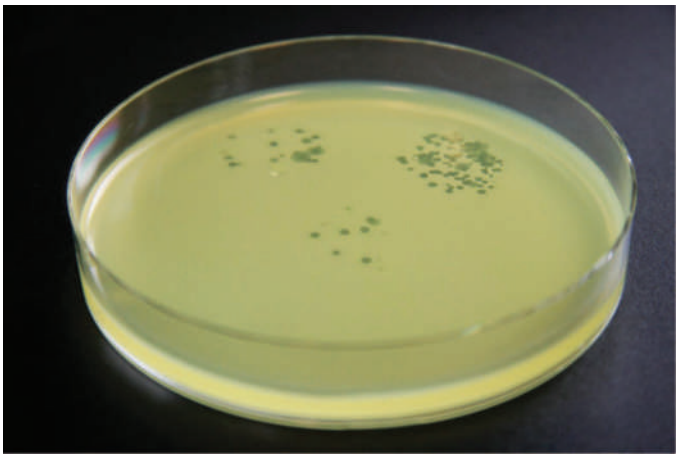


Figure 4 | Experimental assay to show bacteriophage infecting bacteria. Each plaque, which contains approximately 10^5 bacteriophage particles, indicates that a single phage particle has successfully established an infection in a lawn of *Pseudomonas fluorescens* bacteria. This assay provides a simple means to follow the co-evolution of bacterial resistance and bacteriophage infectivity.

microorganisms from the environment. A great advantage of experimental evolution is that it is driven by natural selection, so the traits required can be artificially selected at the level of the population, and the fittest organisms, and crucially those most resilient to competition from foreign invaders, will dominate. Similarly, experimental evolution can be used to evolve biological control agents, such as highly evolvable bacteriophages that could be used to treat antibiotic-resistant bacterial infections[70].

Given these exciting future prospects, we anticipate that experimental evolution using microorganisms will expand at an even faster rate than in recent years. Darwin would no doubt have been delighted by the insights that experimental evolution has provided into understanding and applying his big idea: evolution by natural selection. ■

1. Darwin, C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (John Murray, 1859).
2. Bell, G. *Selection: The Mechanism of Evolution* (Oxford Univ. Press, 2007).
3. Lenski, R. E., Rose, M. R., Simpson, S. C. & Tadler, S. C. Long term experimental evolution in *Escherichia coli*. 1. Adaptation and divergence during 2000 generations. *Am. Nat.* **138,** 1315–1341 (1991).
This paper provided the first demonstration of long-term evolutionary dynamics in real time.
4. Dallinger, W. On the life-history of a minute septic organism: with an account of experiments made to determine its thermal death point. *Proc. R. Soc. Lond.* **27,** 332–350 (1878).
5. Huxley, J. *Evolution: The Modern Synthesis* (Allen & Unwin, 1942).
6. Elena, S. F. & Lenski, R. E. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Rev. Genet.* **4,** 457–469 (2003).
7. Lenski, R. E. & Travisano, M. Dynamics of adaptation and diversification — a 10,000-generation experiment with bacterial populations. *Proc. Natl Acad. Sci. USA* **91,** 6808–6814 (1994).
8. Woods, R., Schneider, D., Winkworth, C. L., Riley, M. A. & Lenski, R. E. Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **103,** 9107–9112 (2006).
9. Ostrowski, E. A., Woods, R. J. & Lenski, R. E. The genetic basis of parallel and divergent phenotypic responses in evolving populations of *Escherichia coli*. *Proc. R. Soc. Lond. B* **275,** 277–284 (2008).
10. Blount, Z. D. & Grogan, D. W. New insertion sequences of *Sulfolobus*: functional properties and implications for genome evolution in hyperthermophilic Archaea. *Mol. Microbiol.* **55,** 312–325 (2005).
11. Goddard, M. R., Godfray, H. C. J. & Burt, A. Sex increases the efficacy of natural selection in experimental yeast populations. *Nature* **434,** 636–640 (2005).
12. Colegrave, N. Sex releases the speed limit on evolution. *Nature* **420,** 664–666 (2002).
13. Fisher, R. A. *The Genetical Theory of Natural Selection* (Oxford Univ. Press, 1930).
14. Muller, H. J. Some genetic aspects of sex. *Am. Nat.* **8,** 118–138 (1932).
15. Drake, J. W. Spontaneous mutation. *Annu. Rev. Genet.* **25,** 125–146 (1991).
16. Giraud, A., Radman, M., Matic, I. & Taddei, F. The rise and fall of mutator bacteria. *Curr. Opin. Microbiol.* **4,** 582–585 (2001).
17. Taddei, F. *et al.* Role of mutator alleles in adaptive evolution. *Nature* **387,** 700–702 (1997).
18. Giraud, A. *et al.* Costs and benefits of high mutation rates: adaptive evolution of bacteria in the mouse gut. *Science* **291,** 2606–2608 (2001).
This paper identifies the selective forces acting on the mutation rate of pathogenic bacteria *in vivo*.
19. Sniegowski, P. D., Gerrish, P. J. & Lenski, R. E. Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* **387,** 703–705 (1997).
20. Pal, C., Macia, M. D., Oliver, A., Schachar, I. & Buckling, A. Coevolution with viruses drives the evolution of bacterial mutation rates. *Nature* **450,** 1079–1081 (2007).
21. de Visser, J., Zeyl, C. W., Gerrish, P. J., Blanchard, J. L. & Lenski, R. E. Diminishing returns from mutation supply rate in asexual populations. *Science* **283,** 404–406 (1999).
22. Gause, G. F. *The Struggle for Existence* (Williams & Wilkins, 1934).
23. Rainey, P. B. & Travisano, M. Adaptive radiation in a heterogeneous environment. *Nature* **394,** 69–72 (1998).
This paper shows that competition in a spatially variable environment drives the diversification of bacteria into spatial niche specialists.
24. MacLean, R. C. & Bell, G. Experimental adaptive radiation in *Pseudomonas*. *Am. Nat.* **160,** 569–581 (2002).
25. Helling, R. B., Vargas, C. N. & Adams, J. Evolution of *Escherichia coli* during growth in a constant environment. *Genetics* **116,** 349–358 (1987).
26. Rosenzweig, R. F., Sharp, R. R., Treves, D. S. & Adams, J. Microbial evolution in a simple unstructured environment — genetic differentiation in *Escherichia coli*. *Genetics* **137,** 903–917 (1994).
27. Treves, D. S., Manning, S. & Adams, J. Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of *Escherichia coli*. *Mol. Biol. Evol.* **15,** 789–797 (1998).
28. Elena, S. F. & Lenski, R. E. Long-term experimental evolution in *Escherichia coli*. VII. Mechanisms maintaining genetic variability within populations. *Evolution* **51,** 1058–1067 (1997).
29. Rozen, D. E. & Lenski, R. E. Long-term experimental evolution in *Escherichia coli*. VIII. Dynamics of a balanced polymorphism. *Am. Nat.* **155,** 24–35 (2000).
30. Novak, M., Pfeiffer, T., Lenski, R. E., Sauer, U. & Bonhoeffer, S. Experimental tests for an evolutionary trade-off between growth rate and yield in *E. coli*. *Am. Nat.* **168,** 242–251 (2006).
31. West, S. A., Diggle, S. P., Buckling, A., Gardner, A. & Griffins, A. S. The social lives of microbes. *Annu. Rev. Ecol. Evol. Syst.* **38,** 53–77 (2007).
32. Crespi, B. J. The evolution of social behavior in microorganisms. *Trends Ecol. Evol.* **16,** 178–183 (2001).
33. Darwin, C. *The Descent of Man, and Selection in Relation to Sex* (John Murray, 1871).
34. Hamilton, W. D. The genetical evolution of social behaviour, I & II. *J. Theor. Biol.* **7,** 1–52 (1964).
35. Harrison, E. F. & Buckling, A. Hypermutability impedes cooperation in pathogenic bacteria. *Curr. Biol.* **15,** 1968–1971 (2005).
36. Griffin, A. S., West, S. A. & Buckling, A. Cooperation and competition in pathogenic bacteria. *Nature* **430,** 1024–1027 (2004).
This paper disentangles the role of relatedness and kin competition in driving the evolution of cooperation.
37. Mehdiabadi, N. J. *et al.* Kin preference in a social microbe. *Nature* **442,** 881–882 (2006).
38. Thompson, J. N. *The Coevolutionary Process* (Univ. Chicago Press, 1994).
39. Hamilton, W. D., Axelrod, R. & Tanese, R. Sexual reproduction as an adaptation to resist parasites. *Proc. Natl Acad. Sci. USA* **87,** 3566–3573 (1990).
40. Bohannan, B. J. M. & Lenski, R. E. Linking genetic change to community evolution: insights from studies of bacteria and bacteriophage. *Ecol. Lett.* **3,** 362–377 (2000).
41. Buckling, A. & Rainey, P. B. Antagonistic coevolution between a bacterium and a bacteriophage. *Proc. R. Soc. Lond. B* **269,** 931–936 (2002).
42. Forde, S. E., Thompson, J. N. & Bohannan, B. J. M. Gene flow reverses an adaptive cline in a coevolving host–parasitoid interaction. *Am. Nat.* **169,** 794–801 (2007).
43. Chao, L., Levin, B. R. & Stewart, F. M. A complex community in a simple habitat: an experimental study with bacteria and phage. *Ecology* **58,** 369–378 (1977).
This paper demonstrates extremely rapid real-time co-evolution between natural enemies, suggesting that co-evolution is a crucial process in ecology and evolution.
44. Buckling, A. & Rainey, P. B. The role of parasites in sympatric and allopatric diversification. *Nature* **420,** 496–499 (2002).
45. Modi, R. I. & Adams, J. Coevolution in bacterial-plasmid populations. *Evolution* **45,** 656–667 (1991).
46. Gratten, J. *et al.* A localized negative genetic correlation constrains microevolution of coat color in wild sheep. *Science* **319,** 318–320 (2008).
47. Benmayor, R., Buckling, A., Bonsall, M. B., Brockhurst, M. A. & Hodgson, D. J. The interactive effects of parasites, disturbance, and productivity on experimental adaptive radiations. *Evolution* **62,** 467–477 (2008).
48. Ochman, H., Lawrence, J. G. & Groisman, E. A. Lateral gene transfer and the nature of bacterial innovation. *Nature* **405,** 299–304 (2000).
49. Ackermann, M., Stearns, S. C. & Jenal, U. Senescence in a bacterium with asymmetric division. *Science* **300,** 1920 (2003).
This was the first demonstration that bacteria can age.
50. Stewart, E. J., Madden, R., Paul, G. & Taddei, F. Aging and death in an organism that reproduces by morphologically symmetric division. *PLoS Biol.* **3,** 295–300 (2005).
51. Ackermann, M., Schauerte, A., Stearns, S. C. & Jenal, U. Experimental evolution of aging in a bacterium. *BMC Evol. Biol.* **7,** 126 (2007).
52. Elena, S. F., Cooper, V. S. & Lenski, R. E. Punctuated evolution caused by selection of rare beneficial mutations. *Science* **272,** 1802–1804 (1996).
53. Bantinaki, E. *et al.* Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. III. Mutational origins of wrinkly spreader diversity. *Genetics* **176,** 441–453 (2007).
54. Schluter, D. *The Ecology of Adaptive Radiations* (Oxford Univ. Press, 2000).
55. Dettman, J. R., Sirjusingh, C., Kohn, L. M. & Anderson, J. B. Incipient speciation by divergent adaptation and antagonistic epistasis in yeast. *Nature* **447,** 585–588 (2007).
56. Wagner, A. *Robustness and Evolvability in Living Systems* (Princeton Univ. Press, 2005).
57. Pal, C. *et al.* Chance and necessity in the evolution of minimal metabolic networks. *Nature* **440,** 667–670 (2006).
58. Harrison, R., Papp, B., Pal, C., Oliver, S. G. & Delneri, D. Plasticity of genetic interactions in metabolic networks of yeast. *Proc. Natl Acad. Sci. USA* **104,** 2307–2312 (2007).
59. Mortlock, R. C. E. *Microorganisms as Model Systems for Studying Evolution* (Plenum, 1984).
60. Wichman, H. A., Badgett, M. R., Scott, L. A., Boulianne, C. M. & Bull, J. J. Different trajectories of parallel evolution during viral adaptation. *Science* **285,** 422–424 (1999).
61. Velicer, G. J. *et al.* Comprehensive mutation identification in an evolved bacterial cooperator and its cheating ancestor. *Proc. Natl Acad. Sci. USA* **103,** 8107–8112 (2006).
62. Dunham, M. J. *et al.* Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **99,** 16144–16149 (2002).
63. De Vos, D. *et al.* Study of pyoverdine type and production by *Pseudomonas aeruginosa* isolated from cystic fibrosis patients: prevalence of type II pyoverdine isolates and accumulation of pyoverdine-negative mutations. *Arch. Microbiol.* **175,** 384–388 (2001).
64. Smith, E. E. *et al.* Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. *Proc. Natl Acad. Sci. USA* **103,** 8487–8492 (2006).
65. Meyer, J. M., Neely, A., Stintzi, A., Georges, C. & Holder, I. A. Pyoverdin is essential for virulence of *Pseudomonas aeruginosa*. *Infect. Immun.* **64,** 518–523 (1996).
66. Bjarnasholt, T. & Givskov, M. Quorum-sensing blockade as a strategy for enhancing host defences againts bacterial pathogens. *Phil. Trans. R. Soc. Lond. B* **362,** 1212–1223 (2007).
67. Perron, G. G., Zasloff, M. & Bell, G. Experimental evolution of resistance to an antimicrobial peptide. *Proc. R. Soc. Lond. B* **273,** 251–256 (2006).
68. Bolotin, A. *et al.* Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nature Biotechnol.* **22,** 1554–1558 (2004).
69. Manefield, M., Griffiths, R. I., Leigh, M. B., Fisher, R. & Whiteley, A. S. Functional and compositional comparison of two activated sludge communities remediating coking effluent. *Environ. Microbiol.* **7,** 715–722 (2005).
70. Levin, B. R. & Bull, J. J. Population and evolutionary dynamics of phage therapy. *Nature Rev. Microbiol.* **2,** 166–173 (2004).

Acknowledgements We thank the Royal Society and the Leverhulme Trust for funding.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence should be addressed to A.B. (angus.buckling@zoo.ox.ac.uk).

NATURE|Vol 457|12 February 2009|doi:10.1038/nature07893

Adaptation and diversification on islands

Jonathan B. Losos[1] & Robert E. Ricklefs[2]

Charles Darwin's travels on HMS *Beagle* taught him that islands are an important source of evidence for evolution. Because many islands are young and have relatively few species, evolutionary adaptation and species proliferation are obvious and easy to study. In addition, the geographical isolation of many islands has allowed evolution to take its own course, free of influence from other areas, resulting in unusual faunas and floras, often unlike those found anywhere else. For these reasons, island research provides valuable insights into speciation and adaptive radiation, and into the relative importance of contingency and determinism in evolutionary diversification.

When Charles Darwin embarked on his five-year odyssey on HMS *Beagle*, a Royal Navy surveying ship, he was an inexperienced, 23-year-old gentleman naturalist. He returned as one of the rising scientific stars of his generation. Many of Darwin's experiences on this trip shaped his thinking, but none was more influential than the five weeks he spent in the Galapagos Islands.

Since Darwin's time, research on islands has continued to advance the understanding of the evolutionary process. Indeed, evolutionists have come to regard islands as laboratories of evolution because island geography lays before them the underlying mechanisms of species formation and adaptive radiation. But islands also vary in many ways, and the evolutionary proliferation of life has progressed much further on some islands than on others. The varied outcomes of evolution in island settings can indicate a great deal about how evolutionary and spatial processes have built biological diversity through the formation and differentiation of species.

In this Review, we discuss what Darwin's time in the Galapagos Islands taught him about evolution and explore what islands have shown about evolution since then. In particular, recent work on islands has demonstrated the importance of geographical isolation (allopatry) in the initial stages of species formation, the role of interactions between species in adaptive radiation, and the effects of both historical happenstance and deterministic factors in the outcome of evolution on islands. In addition, in recent years, there has been a widespread move to reconstruct phylogenetic trees for individual islands on the basis of molecular information. These phylogenies describe the history of island biotas and might provide a way to synthesize evolution and biogeography at scales that bridge single remote islands, archipelagoes and continents.

Darwin in the Galapagos Islands

Three of Darwin's observations on the plants and animals of the Galapagos Islands contributed to his realization that species are not immutable, specially created forms. First, he saw that closely related populations on nearby islands vary in appearance, as he explained in *The Voyage of the Beagle*[1]: "My attention was first thoroughly aroused, by comparing together … the mocking-thrushes, when, to my astonishment, I discovered that all those from Charles Island belonged to one species (*Mimus trifasciatus*); all from Albemarle Island to *M. parvulus*; and all from James and Chatham Islands … belonged to *M. melanotis*."

Second, Darwin noted that the inhabitants of an island usually have close affinities with forms on the adjacent continent, rather than with species occupying similar environments elsewhere in the world. This seemed to suggest that they were not created independently[1]: "It is probable that the islands of the Cape de Verd group resemble, in all their physical conditions, far more closely the Galapagos Islands, than these latter physically resemble the coast of America, yet the aboriginal inhabitants of the two groups are totally unlike; those of the Cape de Verd Islands bearing the impress of Africa, as the inhabitants of the Galapagos Archipelago are stamped with that of America."

Third, after returning to England, Darwin was informed by the ornithologist John Gould that many of the bird species in the Galapagos Islands that seemed to belong to different families were, in fact, all related members of a single family, previously unknown to science — and now known as Darwin's finches. Darwin quickly realized the implications of such phenotypic and ecological diversity[1]: "Seeing this gradation and diversity of structure in one small, intimately related group of birds, one might really fancy that from an original paucity of birds in this archipelago, one species has been taken and modified for different ends."

Almost two centuries later, knowledge of the fauna and flora of the Galapagos Islands is more complete. Darwin's observations on mockingbirds, tortoises and finches have been confirmed and greatly amplified, and many of the islands' other taxa have been shown to demonstrate the same patterns of evolutionary descent and diversification (Fig. 1). Moreover, similar examples of evolutionary diversification have been documented on islands throughout the world. Biologists now recognize that it is not the Galapagos Islands in particular, but islands in general, that present a pageant of natural experiments of great value in studying evolutionary processes.

Islands as nature's test tubes

What is so special about islands? Their small size, distinct boundaries, simplified biotas and the abundance and tameness of island inhabitants all make it easier to observe and interpret patterns of evolution. Moreover, groups of islands can function as replicates in which general evolutionary patterns can be distinguished from unique outcomes.

Two additional attributes that make islands lasting focal points for evolutionary studies — their relative youth and geographical isolation — were clearly identified by Alfred Russel Wallace[2], the co-originator of the theory of evolution by natural selection, in his 1881 book *Island Life*[3]. First, many islands are either volcanic in origin or have been completely under water at some point in their history. These islands emerge above the ocean surface as blank slates for colonization and subsequent

[1]Museum of Comparative Zoology, 26 Oxford Street, Harvard University, Cambridge, Massachusetts 02138, USA. [2]Department of Biology, University of Missouri - St. Louis, 8001 Natural Bridge Road, St. Louis, Missouri 63121, USA.

evolutionary diversification, on which the development of ecological and evolutionary systems can be observed from their beginnings. Each island represents a new opportunity for living forms to appear and proliferate. The first colonists, finding untapped resources and lacking the constraints of a resident biota, often diversify in novel directions. This evolutionary idiosyncrasy is enhanced by unbalanced colonization — strong dispersal abilities are not evenly distributed across the ecological spectrum of continental biotas — with the result that some ecological niches on islands are filled by diversification rather than colonization[4].

Most mainland settings, by contrast, are packed with species, their main burst of evolutionary diversification long over. Such filled communities offer relatively few ecological opportunities, and evolutionary diversification tends to produce small variations on already successful adaptive themes. For example, the evolutionary diversification of continental cardueline finches produced species that are ecologically and morphologically similar, essentially size variations on the standard finch-like way of life of cracking seeds with a stout beak (which is not to belittle the peculiar *Loxia* crossbills). By contrast, on the Hawaiian islands, a cardueline radiation descended from an initial colonist has reproduced most of the variation in the entire order of passerine birds: warblers, grosbeaks, tanagers and creepers, as well as a number of forms (such as *Hemignathus*) that lack continental counterparts (Fig. 2).

Second, many islands and archipelagoes are distant from other land masses, and some have been isolated, with low rates of colonization, for long periods. This allows island biotas to diverge along their own evolutionary trajectories, independent of, and unconstrained by, evolutionary events unfolding elsewhere[4,5]. For old or very isolated islands, the particular mixture of evolutionary lineages that were present when an island separated from other areas, or that arrived with the occasional colonization event, is often unique and markedly different from biotas occurring anywhere else in the world.

If the goal is to study the interplay of ecological and evolutionary processes in the generation of biological diversity, islands (and their terrestrial analogues, lakes[6]) frequently offer the best opportunities. A particularly powerful approach is to combine studies of ongoing natural selection and microevolutionary change with phylogenetic analyses of evolutionary patterns in deeper, macroevolutionary, time[7,8], an approach that in some cases can even be experimental[9].

Figure 1 | Evolutionary diversification and adaptive radiation in the Galapagos Islands. Large ground finch, *Geospiza magnirostris* (**a**), sitting on an *Opuntia* cactus. Both *Geospiza* and *Opuntia* have radiated in the Galapagos Islands, as has the snail genus *Bulimulus* (shown is *Bulimulus reibischi*) (**b**), producing phenotypically differentiated species that have adapted to different parts of the environment. Darwin noted the inter-island variation in Galapagos tortoises (*Geochelone nigra*) (**c**) and mockingbirds (*Nesomimus parvulus*) (**d**), but similar variation occurs in many other taxa on these islands, including marine iguanas (*Amblyrhynchus cristatus*, larger lizard) and lava lizards (genus *Microlophus*, smaller lizard on head of iguana) (**e**). (Panels **a**, **d** and **e** courtesy of H. Snell (Visual Escapes); panel **b** courtesy of C. Parent (University of Texas, Austin).)

Evolutionary processes

Evolutionary diversification during adaptive radiation involves two processes: the proliferation of species from an initial ancestor to many descendants, and the adaptation of species to use different parts of the environment. Recent debate has swirled around the order in which these processes occur. Does speciation precede ecological divergence and coexistence, or does divergence drive the speciation process[10]? Island studies have richly enhanced the understanding of both processes and the way they are related.

Speciation

Might history have taken a different course if HMS *Beagle* had sailed a different route? Suppose, for example, that it had stopped not at the Galapagos Islands but at Cocos Island, a small, isolated island north of the Galapagos. As with the Galapagos, most endemic species on Cocos Island are related to species on nearby continental areas[11]. But unlike the Galapagos, little species proliferation (cladogenesis) has occurred *in situ*. For example, the island has just one species of Darwin's finch, *Pinaroloxias inornata*.

It is difficult to say what Darwin would have made of the situation on Cocos Island, but much can be learned from the observation that speciation has occurred on some islands but not on others. There are almost no examples of an ancestral bird species splitting into two descendants on islands smaller than Madagascar[12,13], for example, and the few purported cases are debatable[14,15]. Similarly, *Anolis* lizards have speciated profusely on the larger islands of the Greater Antilles but hardly at all on the Lesser Antilles, even though some of the islands are reasonably large (Guadeloupe has an area of 1,628 km^2), have been occupied by anoles for millions of years, and seemingly provide the range of habitats to which species have adapted in the Greater Antilles[16–18]. Indeed, small islands formerly connected to larger land masses have substantially more species[19]. Speciation in Galapagos snails also occurs only on larger islands, although the area required for *in situ* speciation is smaller than for lizards or birds[20].

Such a lack of speciation on small islands suggests that divergent natural-selection pressures are not sufficient to split one species into two. Although there are exceptions (including the sister taxa of palms, *Howea* spp., on Lord Howe Island[21,22]), the overwhelming absence of species splitting on small islands suggests that allopatry (geographical isolation) is required — a conclusion reached by Ernst Mayr as a result of his studies of island birds[23] — and that a minimum island size exists below which islands lack opportunities for geographical isolation. Alternatively, some proposed mechanisms of non-allopatric speciation (such as speciation across ecological gradients) in theory also require a minimum area (and population size)[24], but these mechanisms seem unlikely for islands as large and environmentally heterogeneous as Guadeloupe.

The importance of allopatry for speciation is also demonstrated by patterns of species generation within archipelagoes. For example, although birds do not diverge into multiple species on small islands, they do so readily on archipelagoes composed of small islands[25]. The evolutionary radiation of Darwin's finches is the best-known example, with 13 species having been produced in the Galapagos archipelago. Detailed examination of variation among populations and microevolutionary changes occurring within populations strongly supports the classic allopatric model of speciation. Populations become isolated on different islands,

and when they become sympatric as a result of one population colonizing an island already occupied by another, either they have already attained species-level distinctiveness or ecological and evolutionary processes acting in sympatry reinforce pre-existing differences, completing the speciation process[8,26] (Fig. 3).

Once islands exceed the speciation threshold, the rate of speciation has been shown to increase with island size, at least in Greater Antillean anoles[16] and Galapagos snails[20,27]. By contrast, island age seems to have a stronger effect than island area in some archipelagoes, with islands of intermediate age having the most species and older islands losing species by extinction as their habitats degrade from erosion and loss of area[28–30].

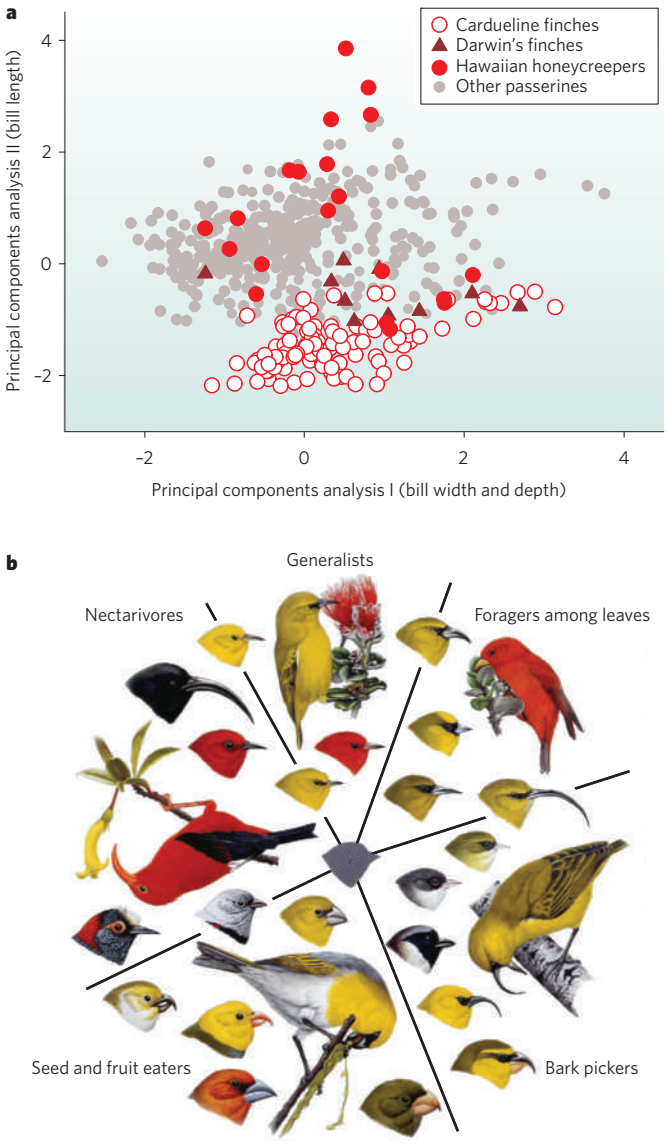


Figure 2 | Distribution of beak shapes in passerine birds, illustrating the tremendous diversification of morphology in Hawaiian honeycreepers. **a**, Graph showing variation in beak morphology in a sample of passerine birds worldwide. Variation is represented by scores on the first two axes of a principal components analysis (which derives uncorrelated axes of variation as linear combinations of the original variables). Towards the bottom right, beaks are short and stout; towards the top left, they are long and slender. Mainland cardueline finches have diversified primarily in bill width and depth, retaining the basic finch-like beak shape, whereas the Hawaiian honeycreepers have also diversified in bill length. (Data from ref. 35.) **b**, Adaptive radiation in Hawaiian honeycreepers, showing how beak shapes are adapted for particular food types. (Panel **b** reproduced, with permission, from ref. 74; courtesy of D. Pratt (North Carolina Museum of Natural Sciences, Raleigh).)

In addition, on old islands, the distributions of closely related allopatric species might expand, bringing the species into sympatry and leading to the extinction of one of the species through competition[31].

Two explanations could account for the relationship between island size and prevalence of speciation. One possibility is that larger islands may present more opportunities for allopatric isolation thanks to their higher elevation and greater topographical and ecological complexity[32,33], as well as a greater opportunity for fragmentation by high sea levels or other geological events[34]. Alternatively, island size is often correlated with ecological diversity, so larger islands might have more niche space and therefore allow the coexistence of more species, even if rates of species generation do not change with area[35]. The evidence on Galapagos snails supports the latter possibility: vegetation diversity, which is an index of niche availability, predicts the number of within-island speciation events better than island area[20,27].

The extent of species diversification on islands is also affected by the degree of isolation. Gene flow between close islands prevents the divergence of populations. At the opposite extreme, populations on widely separated islands differentiate readily but rarely colonize new islands to build up local species numbers. So diversification within archipelagoes requires the appropriate correspondence of geography and dispersal ability[36,37].

Adaptive radiation

The evolutionary exuberance of some island clades is impressive. Every naturalist has a favourite example, perhaps the 30-odd species of silversword plant (Asteraceae) that occupy almost all terrestrial habitats in the Hawaiian islands and exhibit a vast range of morphologies, including trees, erect and compact shrubs, lianas, and branched and unbranched rosettes. Or perhaps it is the roughly 500 species of Hawaiian *Drosophila*, including species that occupy a wide variety of habitats and display a range of phenotypic variation far outstripping that of other *Drosophila*, including differences in body size, wing, leg, antennal and mouthpart morphology, and head shape[38,39].

Adaptive radiation is the outcome of speciation and adaptation in the context of ecological opportunity. It begins with the colonization of a species-poor environment. Allopatric speciation and subsequent recolonization of ancestral islands or areas within islands leads to the sympatry of two or more species. The abundant populations of these species now compete for resources, with selection favouring adaptations that reduce competition between species, including morphological divergence and resource specialization. Repeated again and again, this sequence of species production and character displacement in sympatry can produce a clade of endemic island taxa that are adapted and specialized to use a broad spectrum of ecological space. An unresolved question concerns the extent to which prior ecological divergence in allopatry is necessary to allow coexistence in sympatry, after which evolutionary divergence is driven by character displacement[8,40].

As recently as the 1980s, however, many ecologists questioned the evolutionary significance of character displacement[41]. Most now accept its role in diversification, and some of the best examples come from island settings[10,42]. Islands provide unique opportunities to study character displacement because pairs of species often occur in sympatry on some islands and alone on others. The classic signature of character displacement is that species are more dissimilar in sympatry than in allopatry, although additional genetic, phylogenetic and functional data are also needed[10]. In a recent study on Darwin's finches[43], Peter and Rosemary Grant showed that natural selection in response to drought conditions favoured increased beak size (which is associated with eating larger seeds) in *Geospiza fortis* in the absence of a larger competing species, but decreased beak size (which is more suitable for gleaning smaller seeds) in the presence of such competition.

Adaptive radiation can also be brought about in other ways. For example, plant species that colonize different climatic zones might diverge physiologically and morphologically as they adapt to diverse biotic environments[39,44–46]. Secondary sympatry of such forms could result if their differences allowed them to come back into contact and

coexist in different microclimatic niches. In addition, some adaptive radiations have produced species on multiple trophic levels — with some species eating others — suggesting that predator–prey co-evolution, as well as competitive interactions, can have a role in driving adaptive radiation[10].

Clades of phenotypically disparate species are also often called adaptive radiations, but other processes, such as sexual selection (prominent in Hawaiian crickets[32]) or even founder effects and genetic drift[23,47], can shape phenotypic divergence within a clade[48]. In some cases — including radiations of birds that have diversified in beak size and shape, and radiations of lizards that have diverged in limb and toe-pad dimensions[8,18] — the functional and ecological significance of phenotypic diversity is well understood and the adaptive basis of radiation is well established. In other cases, however, phenotypic divergence is presumed to have an underlying adaptive basis without there being supporting evidence, which ideally would come from an integration of phylogenetic, ecological and functional studies[10].

Contingency and determinism

Stephen Jay Gould famously made the claim that if the tape of the history of life could be re-run, from the same starting point and in an identical environmental setting, there would be a different outcome every time[49]. He argued that the contingencies of history are so great that evolutionary diversification almost always takes a unique and unpredictable course.

Independent radiations of species on different islands or island groups are not strict tests of Stephen Jay Gould's postulate because of the added level of contingency, notably variations in the ecological setting and in the initial colonists. This accidental happenstance of colonization may greatly influence the course of subsequent evolutionary diversification. Consider, for example, evolutionary diversification on the islands of New Zealand, which has occurred in the almost complete absence of mammals. As a result, the fauna of pre-human New Zealand was unlike that of anywhere else[50,51] and included the following: a radiation of moas (flightless browsing ratites up to 3 m in height); a flightless, nocturnal, herbivorous parrot; a carrion-eating parrot; kiwis; enormous raptors; giant orthopterans and weevils; copepods that lived on the forest floor; and primarily terrestrial bats that "represent the bat family's attempt to produce a mouse"[52].

A similar story played out in Madagascar, which became isolated from the Gondwanan land mass during the Late Cretaceous period. Although Madagascar is close to Africa, many groups of birds and mammals that subsequently arose on the African continent failed to cross the gap. In their absence, many other groups took their place, including the following: radiations of herbivorous elephant birds (flightless ratites even larger than moas); lemurs ranging in size from 30 g to 200 kg, including recently extinct forms potentially convergent to tree sloths, ground sloths, gorillas and koalas; giant tortoises; chameleons, which originated in Madagascar; a radiation of carnivores related to mongooses, including one species the same size and shape as a mountain lion; hedgehog look-alikes; and radiations of birds that were originally placed in a multitude of different families[5,53]. On smaller islands, the contingencies of history and opportunism of natural selection are amply demonstrated by the unusual adaptations seen in species such as the carnivorous and aquatic caterpillars of the Hawaiian islands[54] and the frugivorous, metre-long, prehensile-tailed skink (*Corucia zebrata*) of the Solomon Islands.

Nonetheless, many island radiations have produced species resembling forms that evolved independently elsewhere; that is, evolution had similar outcomes, disparate origins notwithstanding[4,5]. A favourite pastime of evolutionary biologists is to take an island radiation, such as Darwin's finches or Malagasy vangas, and, for each of its species, identify a continental ecological analogue, each usually from a different family. However, it is important to remember that cases such as this do not count as identical outcomes of the Gouldian evolutionary replay, both because some island species have no counterparts elsewhere and because the comparisons are not with the fauna of any particular locality but with species found all over the world.

A strict definition of evolutionary determinism in adaptive radiations would include species-for-species matching[55] between independently evolving clades. Such precise evolutionary convergence is rare[56,57] and is only accomplished among closely related clades diversifying in the same region, for which the starting conditions may be similar, satisfying the setting for Stephen Jay Gould's parable, if not its outcome. Probably the most thoroughly documented example is the radiation of *Anolis* lizards on islands in the Greater Antilles. On each island, evolutionary diversification has proceeded for the most part independently, producing on each

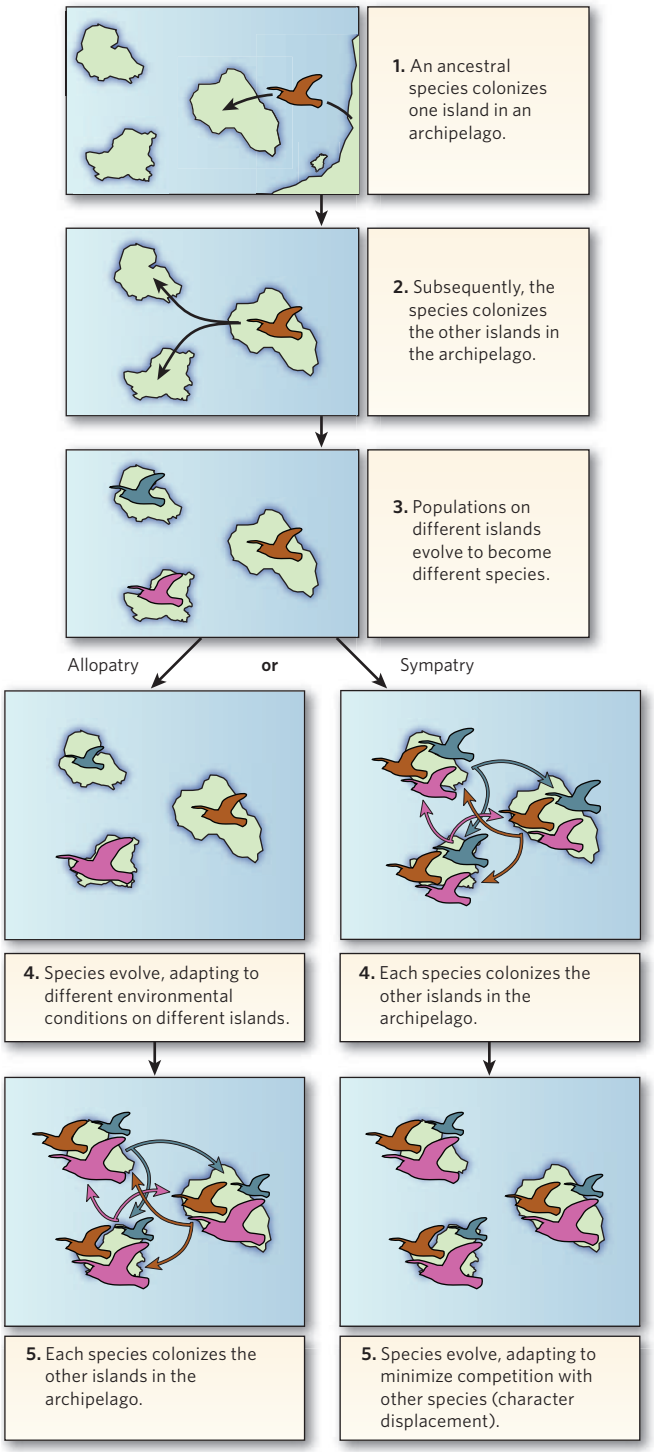


Figure 3 | Two variants of a model of allopatric speciation and subsequent sympatry in an archipelago. The top three panels apply to both models. The lower panels illustrate the possible roles of ecological divergence in allopatry (left) and character displacement in sympatry (right). (Conceptual framework from refs 8 and 16.)

island the same set of habitat specialists adapted to use different parts of the vegetation[18,58,59] (Fig. 4). A similar phenomenon is seen among land snails of the Bonin Islands, near Japan, in which the same set of morphologically convergent habitat specialists has evolved independently on several islands[60].

Simon Conway Morris, responding to Stephen Jay Gould, argued that identical phenotypic outcomes should not be expected; instead, convergence at the level of functional niche filling should be expected, even if the way in which the niches are filled differs[61]. For example, in the absence of woodpeckers — birds that are specialized to extract grubs and other larvae from within woody surfaces but which disperse poorly — island species have evolved a remarkable variety of adaptations to accomplish the same end. These include the Malagasy aye-aye (*Daubentonia madagascariensis*), a lemur that uses its elongated finger to probe holes in wood; the tool-using woodpecker finch of the Galapagos (*Camarhynchus pallidus*), which probes holes with a cactus spine; the New Zealand huia (*Heteralocha acutirostris*), sadly now extinct, for which the short-billed male chiselled into wood to capture insects, and the female (which had a narrower but longer curved bill) reached into crevices inaccessible to the male to extract prey, essentially dividing between the sexes the two functions served by the woodpecker's stout bill and long, extensible tongue; and the Hawaiian akiapolaau (*Hemignathus munroi*), which excavates holes with its short lower mandible and then extracts prey with its long and decurved upper mandible.

Deterministic patterns of evolution can also be seen in the repeated trends observed across many islands. For example, Darwin noted that most tree species make poor long-distance colonizers because of the large size of their seeds[62]. As a result, herbaceous species colonizing new islands often find themselves in open, treeless settings, where selection favours increased size to compete effectively for light. This leads to the evolution of tree-like morphologies in plant clades that never produce such phenotypes in mainland settings[4,63].

Another example is the 'island rule,' the tendency — contested by some[64] — for small mammals to become larger and for large mammals to become smaller[65]. Island dwarfism has received the most attention, owing to the evolution on islands around the world of miniature elephants, the smallest of which was only 1 m high at the shoulder, and of hippopotamuses the size of pigs on Mediterranean islands. Interest in dwarfism was heightened by the discovery of a putative new species of small hominid that until recently lived on the island of Flores, in Indonesia[66]. Less marked increases

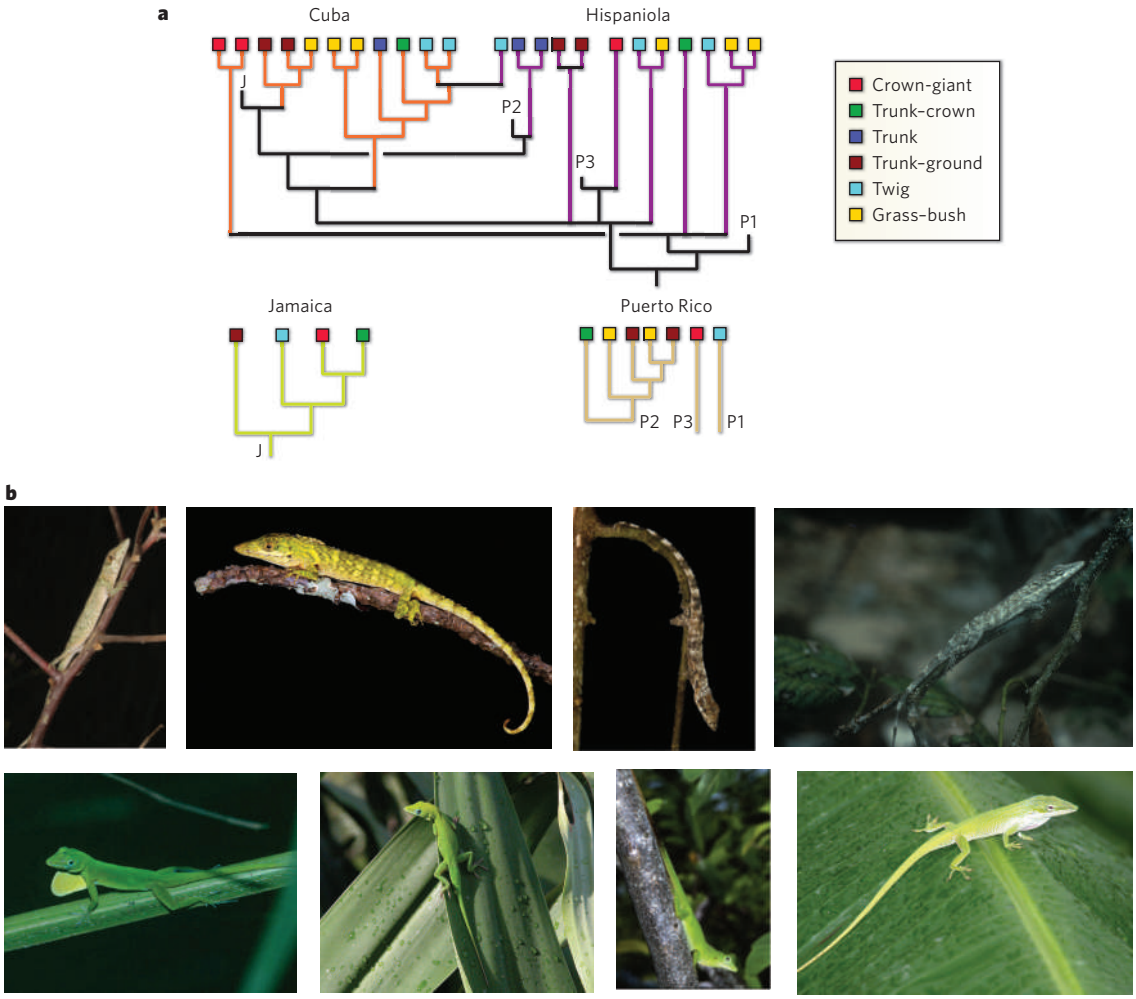


Figure 4 | Independent evolution of a set of *Anolis* lizard habitat specialists in the Greater Antilles. a, Phylogenetic tree. Horizontal black lines connect the clades on different islands and represent either overwater dispersal or ancient geological connections among islands[18]. Coloured lines represent the four locations. For ease of presentation, the Jamaican clade (J) and Puerto Rican clades (P1, P2 and P3) are shown separately but fit into the phylogeny at the positions marked. (Data are based on the molecular phylogeny in ref. 75.) **b**, Images show twig anoles and trunk–crown anoles. Members of each class on different islands of the Greater Antilles are not closely related to each other. Top row shows twig anoles — from left to right: *Anolis occultus*, Puerto Rico; *Anolis insolitus*, Hispaniola; *Anolis valencienni*, Jamaica; *Anolis angusticeps*, Cuba. Bottom row shows trunk–crown anoles — from left to right: *Anolis evermanni*, Puerto Rico; *Anolis chlorocyanus*, Hispaniola; *Anolis grahami*, Jamaica; *Anolis porcatus*, Cuba. (Images courtesy of the following individuals. Top row, left to right: A. Sanchez, M. Landestoy, L. Mahler (Harvard University, Cambridge, Massachusetts) and J. Losos. Bottom row, left to right: J. Losos, M. Losos, K. de Queiroz (National Museum of Natural History, Smithsonian Institution, Washington DC) and A. Sanchez.)

in body size, termed gigantism, are seen in many insular rodents and other small mammals. Similar evolutionary shifts occur in other groups of vertebrates[65], although some truly gigantic forms, such as moas and elephant birds, defy the rule.

Explanations for the island rule are still debated. Reduced predation on islands might release animals from being either very large, to defend themselves against predators, or very small, to hide from predators[65]. Intermediate body size also might be more favourable energetically, allowing the maximum allocation of energy to growth and reproduction[67,68]. According to this theory, selection would push mammals in the direction of the optimum, intermediate, size, but this would be countered by competitive and predatory pressures that are stronger in more species-rich mainland settings.

Other examples of deterministic patterns on islands are the evolution of flightlessness in many birds and insects, the repeated loss of dispersal abilities in both plants and animals, and the evolution of tameness in many vertebrates[4,63] (all phenomena that were documented by Darwin[1,62]).

Paradoxically, islands provide excellent examples of both contingency and determinism in evolution. The contingencies of lineage distribution and colonization success have seeded islands with markedly different starting points; the result is that many island faunas are unique. Yet determinism is seen at several levels, from the existence of general evolutionary trends to the evolutionary filling of the same functional roles and the occasional precise replication of communities of specialists.

Post-Darwinian islands

Many of the principles that form the modern theory of evolution, including divergence of populations in allopatry, evolutionary modification of form and function, and the diversification that occurs during adaptive radiations, were foreshadowed by Darwin thanks to his acquaintance with islands[1,62]. Subsequent research on island biotas has contributed to the understanding of evolutionary and ecological systems in ways that were not envisaged by Darwin. Here we mention two of these: the geographical and evolutionary dynamics of island biotas, and extinction.

One of Darwin's major insights was that the flora and fauna of the Galapagos Islands were ultimately derived from elsewhere. But another century would pass before Edward O. Wilson[69,70] inferred the dynamic interaction of dispersal and evolution from the distributions of ant taxa throughout Melanesia. Wilson realized that taxa went through cycles of dispersal to distant islands from coastal habitats, followed by adaptation to forest interior environments, loss of dispersal ability and eventual extinction. Because species might evolve secondary coastal distributions and initiate new cycles of expansion to more distant islands, Wilson called this pattern the taxon cycle. Molecular phylogenetic analyses of birds in the West Indies have provided the best evidence for taxon cycles, confirmed the temporal sequence of the distribution patterns that represent expansion and contraction phases, and strongly suggested that the cycles are generated by co-evolutionary relationships between species and their predators and pathogens[71].

Darwin knew that many fossils represented extinct organisms, and he used extinction to explain the absence of intermediate forms. But direct observation was impossible, and the causes of natural extinctions (beyond major catastrophes) remain largely out of reach even now. Extinction is the converse of speciation and diversification, but its potential role as a selective agent and potentially creative force has not been resolved. Robert MacArthur and Wilson[72] suggested that the number of species on an island achieves a steady state, with species that go extinct being replaced by colonists from elsewhere. They suggested that the rate of extinction is inversely related to island size, but neither the immediate causes of extinction nor the cause-and-effect relationship between colonization and extinction have been worked out. Islands, and particularly archipelagoes, seem to be an ideal setting for investigating extinction because of the discrete nature of island populations and, for many strongly dispersing taxa, the inference that gaps in distribution represent extinction events. The phylogenetic studies of birds in the West Indies suggested that the probability of extinction increases with the age of an island population and, as MacArthur and Wilson supposed, is inversely related to island area[73].

Darwin's crucial insight was to recognize the connection between evolution and geography — that isolated lineages can evolve independently, ultimately forming new species. Although Darwin was able to infer this principle from observations of close relatives on the different islands of the Galapagos archipelago, biologists now have the tools to examine the evolutionary process more directly. Detailed population-level studies can now chart the course of evolution over short time periods, directly measuring natural selection and examining the extent to which its strength and direction change over time. In turn, phylogenetic techniques can demonstrate the histories of island biotas. These histories now allow the study of trait evolution, from which a general theory of adaptive radiation is beginning to emerge. By integrating these two approaches, biologists have the opportunity to connect pattern and process to test Darwin's postulate that natural selection is the primary engine driving evolutionary change.

One might wonder what Darwin could have accomplished with today's more complete biogeographical and palaeontological information and knowledge of evolutionary mechanisms and phylogenetic relationships. All we can say for sure is that most of what is known today about evolution and the diversification of life is the direct result of his insights. Clearly, Darwin was the right man for the time, and the fruits of his brief visit to the Galapagos Islands in 1835 remain with us. ■

1. Darwin, C. *Journal of Researches into the Natural History and Geology of the Countries Visited During the Voyage of H.M.S.* Beagle *Round the World, under the Command of Capt. Fitz Roy, R.N.* 2nd edn (John Murray, 1845).
2. Berry, A. & Browne, J. The other beetle-hunter. *Nature* **453,** 1188–1190 (2008).
3. Wallace, A. R. *Island Life: Or, The Phenomena and Causes of Insular Faunas and Floras, Including a Revision and Attempted Solution of the Problem of Geological Climates* (Harper, 1881).
4. Carlquist, S. *Island Biology* (Columbia Univ. Press, 1974).
5. Leigh, E. G. Jr, Hladik, A., Hladik, M.-C. & Jolly, A. The biogeography of large islands, or how does the size of the ecological theater affect the evolutionary play? *Rev. Ecol.* **62,** 105–168 (2007).
6. Fryer, G. Endemism, speciation and adaptive radiation in great lakes. *Environ. Biol. Fish.* **45,** 109–131 (1996).
7. Grant, P. R. & Grant, B. R. Unpredictable evolution in a 30-year study of Darwin's finches. *Science* **296,** 707–711 (2002).
8. Grant, P. R. & Grant, B. R. *How and Why Species Multiply: The Radiation of Darwin's Finches* (Princeton Univ. Press, 2008).
This book provides a comprehensive and up-to-date review of the evolutionary diversification of Darwin's finches.
9. Losos, J. B., Schoener, T. W., Langerhans, R. B. & Spiller, D. A. Rapid temporal reversal in predator-driven natural selection. *Science* **314,** 1111 (2006).
10. Schluter, D. *The Ecology of Adaptive Radiation* (Oxford Univ. Press, 2000).
This book gives an insightful and innovative treatment of adaptive radiation, with discussion of, and examples from, island radiations.
11. Kirkendall, L. R. & Jordal, B. H. The bark and ambrosia beetles (Curculionidae, Scolytinae) of Cocos Island, Costa Rica and the rule of mating systems in island zoogeography. *Biol. J. Linn. Soc.* **89,** 729–743 (2006).
12. Diamond, J. M. Continental and insular speciation in Pacific land birds. *Syst. Zool.* **26,** 263–268 (1977).
13. Coyne, J. A. & Price, T. D. Little evidence for sympatric speciation in island birds. *Evolution* **54,** 2166–2171 (2000).
This paper demonstrates that there is lack of speciation on small islands and suggests that non-allopatric speciation does not occur in birds.
14. Ryan, P. G., Bloomer, P., Moloney, C. L., Grant, T. J. & Delport, W. Ecological speciation in south Atlantic island finches. *Science* **315,** 1420–1423 (2007).
15. Grant, P. R. & Grant, B. R. in *The Theory of Island Biogeography Revisited* (eds Losos, J. B. & Ricklefs, R. E.) (Princeton Univ. Press, in the press).
16. Losos, J. B. & Schluter, D. Analysis of an evolutionary species-area relationship. *Nature* **408,** 847–850 (2000).
17. Thorpe, R. S., Surget-Groba, Y. & Johansson, H. The relative importance of ecology and geographic isolation for speciation in anoles. *Phil. Trans. R. Soc. Lond. B* **363,** 3071–3081 (2008).
18. Losos, J. B. *Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles* (Univ. California Press, 2009).
19. Rand, A. S. Competitive exclusion among anoles (Sauria: Iguanidae) on small islands in the West Indies. *Breviora* **319,** 1–16 (1969).
20. Losos, J. B. & Parent, C. E. in *The Theory of Island Biogeography Revisited* (eds Losos, J. B. & Ricklefs, R. E.) (Princeton Univ. Press, in the press).
21. Savolainen, V. *et al.* Sympatric speciation in palms on an oceanic island. *Nature* **441,** 210–213 (2006).
22. Gavrilets, S. & Vose, A. Case studies and mathematical models of ecological speciation. 2. Palms on an oceanic island. *Mol. Ecol.* **16,** 2910–2921 (2007).
23. Mayr, E. *Animal Species and Evolution* (Belknap Press, 1963).
24. Gavrilets, S. & Vose, A. Dynamic patterns of adaptive radiation. *Proc. Natl Acad. Sci. USA* **102,** 18040–18045 (2005).
25. Mayr, E. & Diamond, J. *The Birds of Northern Melanesia: Speciation, Ecology, and Biogeography* (Oxford Univ. Press, 2001).

26. Lack, D. *Darwin's Finches* (Cambridge Univ. Press, 1947).
27. Parent, C. E. & Crespi, B. J. Sequential colonization and diversification of Galapagos endemic land snail genus *Bulimulus* (Gastropoda, Stylommatophora). *Evolution* **60,** 2311–2328 (2006).
The phylogenetic analysis reported in this paper indicates that the prevalence of speciation in Galapagos snails increases with island area and vegetation diversity.
28. Emerson, B. C. & Oromí, P. Diversification of the forest beetle genus *Tarphius* on the Canary Islands, and the evolutionary origins of island endemics. *Evolution* **59,** 586–598 (2005).
29. Whittaker, R. J., Triantis, K. A. & Ladle, R. J. A general dynamic theory of oceanic island biogeography. *J. Biogeogr.* **35,** 977–994 (2008).
This paper provides an important synthesis of the biological and geological factors that shape island evolution.
30. Gillespie, R. G., Claridge, E. M. & Goodacre, S. L. Biogeography of the fauna of French Polynesia: diversification within and between a series of hot spot archipelagoes. *Phil. Trans. R. Soc. Lond. B* **363,** 3335–3346 (2008).
31. Gillespie, R. Community assembly through adaptive radiation in Hawaiian spiders. *Science* **303,** 356–359 (2004).
32. Shaw, K. L. in *Hawaiian Biogeography. Evolution on a Hot Spot Archipelago* (eds Wagner, W. L. & Funk, V. A.) 39–56 (Smithsonian Institution Press, 1995).
33. Otte, D. in *Speciation and its Consequences* (eds Otte, D. & Endler, J. A.) 482–526 (Sinauer, 1989).
34. Gifford, M. E. & Larson, A. *In situ* genetic differentiation in a Hispaniolan lizard (*Ameiva chrysolaema*): a multilocus perspective. *Mol. Phylogenet. Evol.* **49,** 277–291 (2008).
35. Lovette, I. J., Bermingham, E. & Ricklefs, R. E. Clade-specific morphological diversification and adaptive radiation in Hawaiian songbirds. *Proc. R. Soc. Lond. B* **269,** 37–42 (2002).
36. Heaney, L. R. Dynamic disequilibrium: a long-term, large-scale perspective on the equilibrium model of island biogeography. *Glob. Ecol. Biogeogr.* **9,** 59–74 (2000).
37. Parent, C. E., Caccone, A. & Petren, K. Colonization and diversification of Galapagos terrestrial fauna: a phylogenetic and biogeographical synthesis. *Phil. Trans. R. Soc. Lond. B* **363,** 3347–3361 (2008).
38. Kambysellis, M. P. & Craddock, E. M. in *Molecular Evolution and Adaptive Radiation* (eds Givnish, T. J. & Sytsma, K. J.) 475–509 (Cambridge Univ. Press, 1997).
39. Carlquist, S., Baldwin, B. G. & Carr. G. *Tarweeds and Silverswords: Evolution of the Madiinae* (Missouri Botanical Garden Press, 2003).
40. Slatkin, M. Ecological character displacement. *Ecology* **71,** 163–177 (1980).
41. Simberloff, D. & Boecklen, W. Santa Rosalia reconsidered: size ratios and competition. *Evolution* **35,** 1206–1228 (1981).
42. Dayan, T. & Simberloff, D. Ecological and community-wide character displacement: the next generation. *Ecol. Lett.* **8,** 875–894 (2005).
43. Grant, P. R. & Grant, B. R. Evolution of character displacement in Darwin's finches. *Science* **313,** 224–226 (2006).
This paper provides perhaps the best-documented example of character displacement and shows how the presence of a congeneric species changes the course of adaptation.
44. Francisco-Ortega, J., Jansen, R. K. & Santos-Guerra, A. Chloroplast DNA evidence of colonization, adaptive radiation, and hybridization of the Macaronesian fauna. *Proc. Natl Acad. Sci. USA* **93,** 4085–4090 (1996).
45. Givnish, T. J., Montgomery, R. A. & Goldsteing, G. Adaptive radiation of photosynthetic physiology in the Hawaiian lobeliads: light regimes, static light responses, and whole-plant compensation points. *Am. J. Bot.* **91,** 228–246 (2004).
46. Meimberg, H. *et al.* Molecular evidence for adaptive radiation of *Micromeria* Benth. (Lamiaceae) on the Canary Islands as inferred from chloroplast and nuclear DNA sequences and ISSR fingerprint data. *Mol. Phylogenet. Evol.* **41,** 566–578 (2006).
47. Carson, H. L. & Templeton, A. R. Genetic revolutions in relation to speciation phenomena: the founding of new populations. *Annu. Rev. Ecol. Syst.* **15,** 97–131 (1984).
48. Givnish, T. J. in *Molecular Evolution and Adaptive Radiation* (eds Givnish, T. J. & Systma, K. J.) 1–54 (Cambridge Univ. Press, 1997).
49. Gould, S. J. *Wonderful Life* (Norton, 1989).
50. Daugherty, C. H., Gibbs, G. W. & Hitchmough, R. A. Mega-island or micro-continent? New Zealand and its fauna. *Trends Ecol. Evol.* **8,** 437–442 (1993).
51. Goldberg, J., Trewick, S. A. & Paterson, A. M. Evolution of New Zealand's terrestrial fauna: a review of molecular evidence. *Phil. Trans. R. Soc. Lond. B* **363,** 3319–3334 (2008).
This detailed account of diversification shows that most of New Zealand's biota results from overwater colonization, rather than survival since New Zealand separated from Australia.
52. Diamond, J. in *Ecological Restoration of New Zealand Islands* (eds Towns, D. R., Daugherty, C. H. & Atkinson, I. A. E.) 3–8 (Dept Conservation, Wellington, 1990).
53. Schulenberg, T. S. in *The Natural History of Madagascar* (eds Goodman, S. M. & Benstead, J. P.) 1130–1134 (Univ. Chicago Press, 2003).
54. Rubinoff, D. Phylogeography and ecology of an endemic radiation of Hawaiian aquatic case-bearing moths (Hyposmocoma: Cosmopterigidae). *Phil. Trans. R. Soc. Lond. B* **363,** 3459–3465 (2008).
55. Schluter, D. Species-for-species matching. *Am. Nat.* **136,** 560–568 (1990).
56. Ricklefs, R. E. & Travis, J. A morphological approach to the study of avian community organization. *Auk* **97,** 321–338 (1980).
57. Wiens, J. A. Ecological similarity of shrub–desert avifaunas of Australia and North America. *Ecology* **72,** 479–495 (1991).
58. Williams, E. E. in *Lizard Ecology: Studies of a Model Organism* (eds Huey, R. B., Pianka, E. R. & Schoener, T. W.) 326–370 (Harvard Univ. Press, 1983).
59. Losos, J. B., Jackman, T. R., Larson, A., de Queiroz, K. & Rodríguez-Schettino, L. Contingency and determinism in replicated adaptive radiations of island lizards. *Science* **279,** 2115–2118 (1998).
60. Chiba, S. Ecological and morphological patterns in communities of land snails of the genus *Mandarina* from the Bonin Islands. *J. Evol. Biol.* **17,** 131–143 (2004).
61. Conway Morris, S. *Life's Solution: Inevitable Humans in a Lonely Universe* (Cambridge Univ. Press, 2003).
62. Darwin, C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle For Life* (John Murray, 1859).
63. Whittaker, R. J. & Fernández-Palacios, J. M. *Island Biogeography: Ecology, Evolution, and Conservation* 2nd edn (Oxford Univ. Press, 2007).
64. Meiri, S., Cooper, N. & Purvis, A. The island rule: made to be broken? *Proc. R. Soc. B* **275,** 141–148 (2008).
65. Lomolino, M. V. Body size evolution in insular vertebrates: generality of the island rule. *J. Biogeogr.* **32,** 1683–1699 (2005).
66. Brown, P. *et al.* A new small-bodied hominin from the Late Pleistocene of Flores, Indonesia. *Nature* **431,** 1055–1061 (2004).
67. Brown, J. H., Marquet, P. A. & Taper, M. L. Evolution of body-size — consequences of an energetic definition of fitness. *Am. Nat.* **142,** 573–584 (1993).
68. Damuth, J. Cope's rule, the island rule and the scaling of mammalian population density. *Nature* **365,** 748–750 (1993).
69. Wilson, E. O. Adaptive shift and dispersal in a tropical ant fauna. *Evolution* **13,** 122–144 (1959).
70. Wilson, E. O. The nature of the taxon cycle in the Melanesian ant fauna. *Am. Nat.* **95,** 169–193 (1961).
71. Ricklefs, R. E. & Bermingham, E. The concept of the taxon cycle in biogeography. *Glob. Ecol. Biogeogr.* **11,** 353–361 (2002).
72. MacArthur, R. H. & Wilson, E. O. *The Theory of Island Biogeography* (Princeton Univ. Press, 1967).
73. Ricklefs, R. E. & Bermingham, E. History and the species-area relationship in Lesser Antillean birds. *Am. Nat.* **163,** 227–239 (2004).
74. Pratt, H. D. *The Hawaiian Honeycreepers: Drepanididae* (Oxford Univ. Press, 2005).
75. Nicholson, K. E. *et al.* Mainland colonization by island lizards. *J. Biogeogr.* **32,** 929–938 (2005).

Acknowledgements For comments on drafts, we thank A. Berry, R. Gillespie, R. Glor, P. Grant, L. Harmon, C. Parent, T. Price, D. Schluter and R. Whittaker. The National Science Foundation, the National Geographic Society and the Smithsonian Institution have supported our research on island faunas. We are also grateful to numerous individuals and governmental institutions throughout the West Indies for logistical and other support in the field.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence should be addressed to J.B.L. (jlosos@oeb.harvard.edu).

Darwin's bridge between microevolution and macroevolution

David N. Reznick[1] & Robert E. Ricklefs[2]

Evolutionary biologists have long sought to understand the relationship between microevolution (adaptation), which can be observed both in nature and in the laboratory, and macroevolution (speciation and the origin of the divisions of the taxonomic hierarchy above the species level, and the development of complex organs), which cannot be witnessed because it occurs over intervals that far exceed the human lifespan. The connection between these processes is also a major source of conflict between science and religious belief. Biologists often forget that Charles Darwin offered a way of resolving this issue, and his proposal is ripe for re-evaluation in the light of recent research.

Charles Darwin and Alfred Russel Wallace based their insight that organisms evolve by natural selection on four principles[1,2]: first, that organisms have "individual variations" that are faithfully transmitted from parent to offspring; second, that all organisms produce more offspring than are required to replace themselves in the next generation; third, that limited resources create a "struggle for existence" that regulates population size, such that most offspring die without reproducing; and fourth, that the individuals that survive and reproduce are, on average, by virtue of their individual variations, better suited to their local environment than those that do not.

Given these four principles, evolution by natural selection (Darwin's 'principle of descent with modification') naturally follows. Such adaptive modifications within populations over time are now referred to as microevolution. Darwin anticipated that microevolution would be a process of continuous and gradual change.

The term macroevolution, by contrast, refers to the origin of new species and divisions of the taxonomic hierarchy above the species level, and also to the origin of complex adaptations, such as the vertebrate eye. Macroevolution posed a problem to Darwin because his principle of descent with modification predicts gradual transitions between small-scale adaptive changes in populations and these larger-scale phenomena, yet there is little evidence for such transitions in nature. Instead, the natural world is often characterized by gaps, or discontinuities. One type of gap relates to the existence of "organs of extreme perfection", such as the eye, or morphological innovations, such as wings, both of which are found fully formed in present-day organisms without leaving evidence of how they evolved. Another category is that species and higher ranks in the taxonomic hierarchy are often separated by gaps without evidence of a transition between them. These discontinuities, plus the discontinuous appearance and disappearance of taxa in the fossil record, form the modern conceptual divide between microevolution and macroevolution (Box 1).

Most evolutionary biologists think that Darwin explained macroevolution simply as microevolution writ large. In fact, Darwin had rather more to say about the relationship between microevolution and macroevolution and invoked additional principles to define it. It is these additional principles that are of interest here because they are often forgotten in discussions of the relationship between microevolution and macroevolution.

The keys to Darwin's thinking about macroevolution are the 'principle of divergence' and extinction. In this Review, we consider how these principles have fared since the publication of *On the Origin of Species*[1], and we discuss whether Darwin's concept of the relationship between microevolution and macroevolution can provide useful insight today. This relationship continues to generate controversy both within the biological sciences and in the confrontation between science and religion. On the 200th anniversary of Darwin's birth, a status report is surely in order.

Divergence and extinction in 1859

Extinction was a hot topic during Darwin's formative years. Many fossils had been found that were not identifiable as living organisms, but it was not until the early nineteenth century that Georges Cuvier argued that these fossils represented organisms that were extinct. A competing hypothesis was that these creatures lived on in the vast (at that time) unexplored regions of the globe. When US President Thomas Jefferson dispatched Meriwether Lewis and William Clark to explore the interior of North America, he expected them to find living mammoths and mastodons, which he knew about from fossils. The French biologist Jean-Baptiste Lamarck championed another alternative: that these fossil organisms, rather than being extinct, had evolved into contemporary forms. Cuvier, who did not believe in evolution, based his thesis of extinction on detailed anatomical comparisons that emphasized differences between fossil and living forms[3].

Darwin embraced Cuvier's explanation and then extended it by proposing that extinction was a by-product of evolution. Accordingly, as organisms evolve under the struggle for existence, some species acquire superior adaptations and exclude other species through competition or exploitation. In this way, extinction reflects the existence of a tipping point in the ongoing struggle for existence. An increase in the severity of any factors that regulate population size can cause populations to decline in abundance and then disappear.

Darwin argued that the struggle for existence was caused by interactions among organisms and was the dominant factor that shaped how organisms evolve. He considered the physical environment to be of minor importance in evolution. Through his experience as a gardener and his visits to the zoological park in London, he observed that organisms from

[1]Department of Biology, University of California, Riverside, California 92521, USA. [2]Department of Biology, University of Missouri - St. Louis, 8001 Natural Bridge Road, St. Louis, Missouri 63121, USA.

Box 1 | A brief survey of macroevolution

An undercurrent of the debate about the mechanisms of macroevolution is whether natural selection (microevolution) is also the cause of macroevolution. Charles Darwin argued that, although natural selection is the sole mechanism that causes evolution, both divergence and extinction shape the larger-scale patterns that emerge from this process.

Macromutations

The initial phase of this debate focused on the expectation that change caused by natural selection will cause continuous variation. It was argued that if natural variation — such as the distinction between species or ranks above species in the taxonomic hierarchy (for example, genera and families) — is discontinuous, then the underlying mechanism that caused that change must also be discontinuous. Some championed macromutation: the origination of a new species as a result of a single large mutation[61].

Punctuated equilibrium and species sorting

Two patterns of evolution revealed in the fossil record have been argued to be inconsistent with natural selection.

First, Stephen Jay Gould and Niles Eldredge observed that the detailed history of individual lineages reveals prolonged intervals with little or no change (equilibrium or stasis) interspersed with intervals of rapid change (punctuations) that are associated with the origin of new species[62]. They proposed that natural selection could fine-tune organisms during periods of stasis but that another mechanism had to account for punctuated change. Second, the fossil record often reveals species sorting, meaning that some lineages rapidly diversify into new species whereas others decline. Gould argued for species selection as the mechanism to explain both phenomena[63]. Species selection treats species as the unit of selection in the same way that natural selection treats individuals as the unit of selection.

Megaevolution and adaptive radiation

George Gaylord Simpson's proposed mechanism of 'megaevolution'[64] was a modern synthesis (1930s–1950s) proposal for how natural selection can combine with other processes to explain species sorting. It stands in opposition to species selection because macroevolution emerges from microevolutionary processes. Simpson combined the idea of key adaptations, or changes that would allow organisms to expand into previously underutilized environments, with Sewall Wright's theoretical models[65] to explain the sudden appearance and expansion of successful lineages.

The key feature of Wright's models is the adaptive landscape. Adaptive peaks are defined by a combination of characters that must appear together to define a well-adapted phenotype. Peaks are separated from one another by 'valleys', or character combinations that result in reduced fitness. Wright's models invoked the combination of natural selection, genetic drift, mutation and migration, in allowing shifts between peaks. Simpson adapted these models to a logical scenario for how a new lineage could be the product of accelerated peak-to-peak evolution. The fact that such evolution is accelerated and happens in a restricted geographical region means that it is unlikely to be seen in the fossil record.

Simpson later developed the concept of adaptive radiation[17], which still stands as a competing explanation for species sorting because such radiations are caused by natural selection and can account for differences among lineages in the rate of diversification (see the section 'Update on divergence').

Heterochrony

Particular types of natural selection and adaptive response have been credited with the potential to cause rapid morphological evolution. Gould[66] championed heterochrony, or changes in the rate of development of one component of an organism relative to others, as a mechanism for the rapid evolution of descendent species that are a mixture of juvenile and adult characters in their ancestors. One hypothesis for the origin of vertebrates is that they are derived from a tunicate-like ancestor that had an actively swimming larval stage with a nerve cord, but then metamorphosed into adults that lost these characters. A common ancestor of the vertebrates could evolve from such a tunicate-like ancestor with the deletion of metamorphosis and the retention of these larval characters into adulthood. The adults of such a descendent species would not be readily identifiable as being closely related to the adult life stage of its ancestor.

a diversity of climates could survive and reproduce perfectly well when transported to England, so their natural distributions were not limited by their climate tolerance. Instead, their distributions were shaped by their ability to disperse and by the presence of other species. For Darwin, the interactions that define the struggle for existence and shape how organisms evolve were diverse, including competition, predation, parasitism, disease and pollination. This same range of interactions can therefore contribute to extinction.

Darwin's principle of divergence derives from what he thought to be one of the most potent components of the struggle for existence. He argued that the strongest interactions would be among individuals within a population or among closely related populations or species, because these organisms have the most similar requirements. Darwin's principle of divergence predicts that the individuals, populations or species most likely to succeed in the struggle are those that differ most from their close relatives in the way they achieve their needs for survival and reproduction.

The principle of divergence has had strong detractors. Ernst Mayr singled it out as a failed theory[4]. His reasoning paralleled his argument that *On the Origin of Species* is not about the origin of species[5]. Darwin saw each species as an arbitrary point on a continuum of populations that are diverging from one another as a consequence of evolution by natural selection. For this reason, he saw the principle of divergence as acting among individuals within a population or among populations or species. He did not distinguish between these levels of interaction. The crucial contribution of Theodosius Dobzhansky[6] and Mayr[5] to the modern synthesis was to recognize that speciation involves both divergence and the origin of discontinuity, or reproductive isolation. Mayr argued that individuals within a population cannot diverge from one another because they are part of an interbreeding gene pool[4].

If it is accepted that reproductive isolation between species is a prerequisite for divergence, then Darwin's principle must be modified so that the initial stages of diversification do not involve interactions between individuals or between closely related populations. This condition does not invalidate Darwin's principle that divergence and extinction are often a consequence of interactions among close relatives; it merely delays the action of the principle until the reproductively isolated descendants of a common ancestral lineage begin to interact.

Darwin illustrated the combined action of his principle of descent with modification, the principle of divergence, and extinction in the only figure in *On the Origin of Species* (Fig. 1). It showed the link between microevolution and macroevolution. Each branch in the evolutionary tree of life is seen as sprouting 'buds' that are emerging varieties or locally distinct populations. These buds are the consequence of the overproduction of offspring whose individual variations allow them to outcompete others or to probe the environment for open ecological space. Most of these buds go extinct, but some persist, becoming modified and improved descendants that tend to drive their close relatives to extinction, or to fill the unoccupied ecological space. According to Darwin, this combination of replacement and divergence causes cladogenesis: the splitting of one ancestral species into more than one descendant. Continued divergence of form and function between genetically isolated species causes the branches of the tree of life to grow ever farther apart, separated from one another by what seem to be unbridgeable gaps. Darwin argued that the processes of diversification and extinction can explain the gaps that are seen among living species. Divergence pushes lineages apart, and extinction erases the bridge that once joined them, creating the appearance of discontinuity.

How has Darwin's proposal for the relationship between microevolution and macroevolution fared since its publication in *On the Origin of Species*? Here we evaluate three conditions necessary for the structural integrity of Darwin's proposed bridge between microevolution and macroevolution. First, some speciation events should be associated with the divergence of ecologically relevant characteristics among descendent lineages; second, at least some extinction events should be attributed to interactions among closely related species; and third, extinction of some lineages should be linked to the diversification of closely related lineages.

Update on divergence

The principle of divergence has never been an explicit subject of analysis, but its imprint can be found in the early development of evolutionary ecology, when it was established that closely related organisms could inhabit the same environment only if they differed in morphology, habitat use or some other characteristics that reduced their ecological similarity[7]. Such observations fostered the idea that competition among close relatives had a significant role in shaping communities[8–10]. Character displacement, or the evolution of increased dissimilarity among species whose geographical ranges overlap[11], provides a natural experimental test of the hypothesis that selection favours individuals in each species that diverge further from the ecological requirements of the other species. Such displacement has now been well characterized in a variety of organisms[12–16].

Darwin's principle of divergence also figures prominently in the process of adaptive radiation, as originally proposed by George Gaylord Simpson[17] and as defined by Dolph Schluter[18] as "the evolution of ecological and phenotypic diversity within a rapidly multiplying lineage. It occurs when a single ancestor diverges into a host of species that use a variety of environments and that differ in traits used to exploit those environments." Adaptive radiations are commonly associated with diversification in sparsely occupied ecological space, for example following the colonization of a remote island, the survival of a mass extinction event, or the crossing of an adaptive barrier to open new evolutionary opportunities. Competition for resources has been shown to have had a dominant role in some adaptive radiations[12,16]. Perhaps the most famous is the evolution of 13 species of geospizine finches (Darwin's finches) within the Galapagos archipelago. The birds' diverse body sizes, beak shapes and diets are all derived from a single common ancestor that colonized an almost bird-free island[12,13]. Diversification within the much larger areas of continents undoubtedly proceeds in the same way for many types of organism.

In the absence of a fossil record, the radiation of any living group of organisms can be visualized by reconstructing the evolutionary relationships among the living species in a lineage to create a phylogeny[19], usually from DNA sequence data. This approach derives from models for lineage diversification that were originally developed in the palaeontological literature[20]. As well as defining the patterns of relatedness among species, phylogenies can yield more specific information about the tempo and possible mode of evolution that underlies contemporary diversity[19,21–24]. The baseline for such analyses is the temporal distribution of branching points in the phylogeny, which allows certain inferences about rates of speciation and extinction in the past[19,25], although there are sources of uncertainty and bias in doing this[26,27].

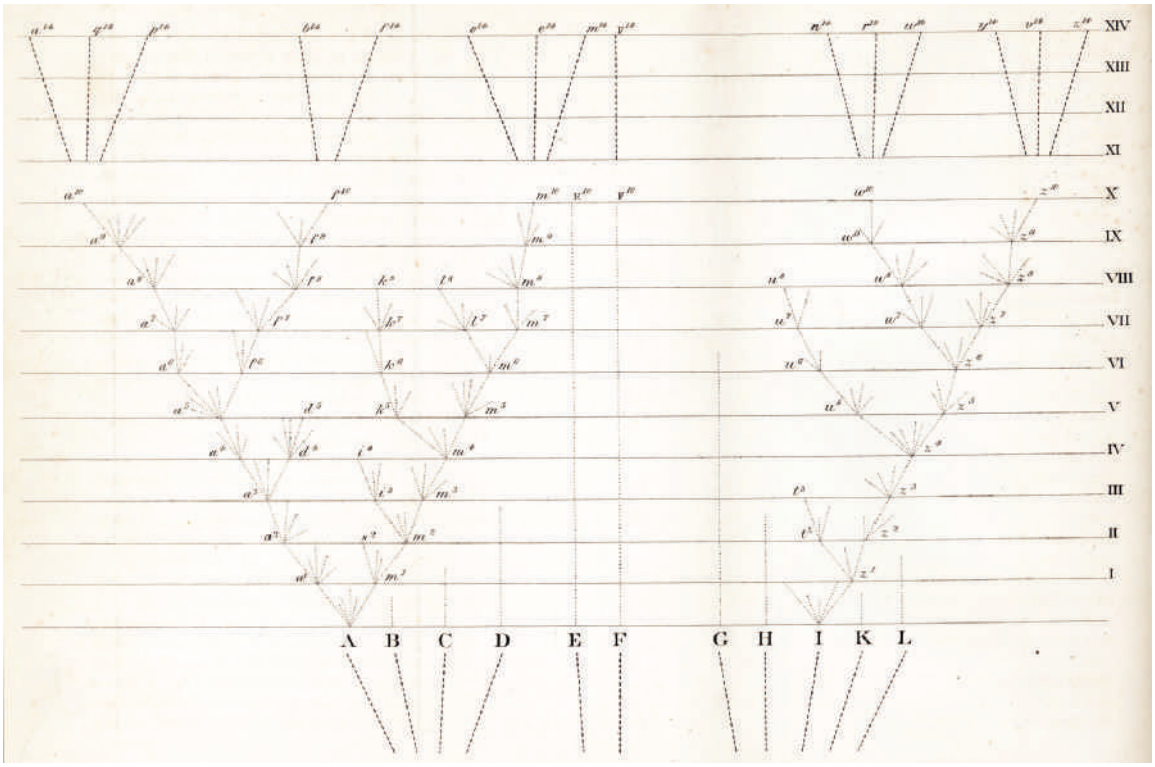


Figure 1 | Darwin's view of the link between microevolution and macroevolution. This figure appears in chapter 4 of Darwin's *On the Origin of Species*. The *x* axis represents a hypothetical ecological variable. The *y* axis represents time. Each horizontal line, associated with roman numerals I to XIV, represents a long but unspecified interval of time. A to L are 11 species in a hypothetical genus. Two of these species (A and I) diversify over time, whereas eight become extinct. One species (F) does not diversify but has surviving descendants, and it represents what Darwin described as "living fossils" — slowly evolving lineages that survived in marginal habitats where they were shielded from interactions with more rapidly diversifying lineages. At each intersection between the diversifying lineages and the divisions in time, the lineage is represented by diverging dashed lines, which are varieties that differ from one another in characteristics and habitat use. Most of these become extinct. Some, labelled with lower-case letters and numerical superscripts, represent distinct descendent subspecies or species. The descendants seen at each time horizon are not simply modified versions of their immediate ancestor but new and improved organisms that outcompeted their parental lineage and drove it to extinction. Thus a^2 is not just a^1 1,000 generations later; it is a daughter lineage that outcompeted a^1. If there is more than one surviving lineage at a node, the survivors tend to be the ones most different from one another. For example, a^1 and m^1 are the most divergent populations derived from A at the end of the first time interval, and these are the ones that survive. As each lineage diversifies, its descendants fan out along the *x* axis, occupying progressively more ecological space. They do so at the expense of the species that lie closest to them on the *x* axis, which become extinct, presumably because they lost out in competition for resources. (Figure reproduced from ref. 1.)

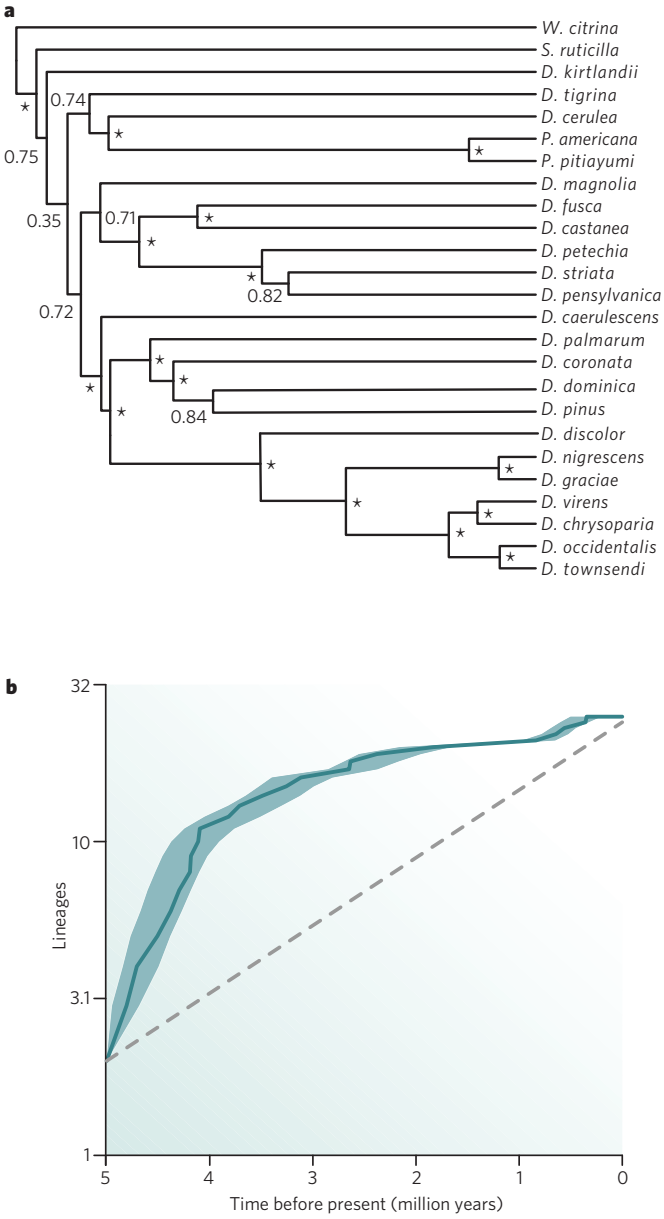


Figure 2 | A plot of lineage through time. This 'lineage through time' (LTT) plot is based on the temporal distribution of branch points in a phylogenetic tree. **a**, Phylogenetic tree for North American wood warblers based on more than 9 kb of mitochondrial and nuclear intron DNA sequence. Species are mainly *Dendroica* spp. but are also from the genera *Parula*, *Wilsonia* and *Setophaga*. The branch lengths are proportional to absolute time. Branch points marked with asterisks are supported by a posterior probability (Bayes' theorem) of >0.95. Numbers at the other branch points are exact probabilities. **b**, The LTT plot derived from this phylogeny. The *y* axis shows the number of lineages; the *x* axis shows the estimated time before present. The solid line represents the absolute value of the log of the number of ancestral lineages (different branches) present at each time interval. The shading is the 95% confidence interval for the number of lineages. The dashed line is what would be seen if the existing species were the product of a constant rate of diversification with no extinction. Note that the rate of accumulation of new lineages is initially high and then levels off. The slope of the lineage-accumulation curve represents the net rate of diversification, which is the rate of formation of new lineages minus the rate of loss of lineages by extinction. Dan Rabosky and Irby Lovette show that this pattern can be explained only by a high initial speciation rate followed by a deceleration in the rate of diversification[60]. This pattern is consistent with density-dependent speciation, or the early diversification of the lineage as it filled the available ecological space before being constrained by resource limitation. (Figure reproduced, with permission, from ref. 30.)

One common pattern is for a lineage to diversify rapidly early in its history, followed by a progressive slowing in the rate of diversification[21,28–30] (Fig. 2). Such a density-dependent cladogenesis[21,24] stems from an initial rapid adaptive radiation that fills the ecological space left open by a mass extinction, for example, or from a spread into a new adaptive zone facilitated by the evolution of some trait that makes the zone more accessible[17,31]. The resultant decline in available ecological space constrains further diversification.

As Darwin envisaged in his principle of divergence, rapid diversification often ensues when a lineage evolves adaptations that enable it to breach an ecological barrier[32]. Examples include the repeated invasion of brackish water environments by the ancestors of the different lineages of mangrove plants[33], the invasion of temperate environments by lineages of trees that evolved freezing tolerance[34–36], and the radiation of skinks that occupied the expanding deserts of Australia as the climate became more arid[37]. Timothy Barraclough and colleagues summarize other radiations that have been associated with the evolution of key innovations[38]. Such radiations also occur on a much grander scale: for example, bats diversified from a single ancestor that evolved flight into the largest order of mammals with more than 1,100 species[39].

Analyses of diversification sometimes reveal other details that are consistent with Darwin's macroevolution theory. For example, the passerine birds include many depauperate lineages that occupy ecologically or geographically marginal habitats[40,41]. These lineages seem to have diversification and extinction rates that are an order of magnitude lower than those of more species-rich clades of passerines, and thus correspond to Darwin's lineage F (Fig. 1), a slowly diversifying lineage isolated from interactions with more rapidly diversifying lineages.

Certainly, Darwin's principle of divergence has been supported by a range of what are now well-characterized and generally accepted evolutionary phenomena.

Update on extinction

Analyses of extinction in the fossil record have been dominated by the discovery of mass extinction events and the later realization that the mass extinctions represent the tail of a distribution of mostly smaller events. Mass extinctions have external causes, including bolide (large crater-forming projectile) impacts, major tectonic events and global climate change[42,43]. These extinctions appear as discrete events involving multiple species and are unrelated to Darwin's proposed mechanism of extinction. Indeed, Darwin's interactive extinctions lie hidden in the background of these events. Modern analyses also almost invariably deal with the appearance and disappearance of genera or families in the fossil record, as species can rarely be distinguished among fossils, although it is the extinction of populations or species that is most relevant to Darwin's principle of divergence. However, even though the fossil record may not be fertile ground for evaluating Darwin's proposed mechanism, species-level originations and extinctions can be identified for some periods and some taxa, such as the recent fossil record for marine bivalves[44], and inferences can be drawn about Darwin's proposed mechanism of extinction.

Analyses of extinction in living organisms fall almost exclusively in the province of conservation biology. Primary causes of extinction in this context include introduced predators and competitors, climate change, and habitat destruction, with the consequent subdivision of once-widespread populations into small isolates. Current research focuses on defining those properties of species that can predict susceptibility to extinction, such as geographical range, population density, life history and trophic level[45–47]. Many of the causes and correlates of contemporary extinctions have the signature of Darwin's emphasis on biotic interactions and the tipping of the balance of factors that regulate population size. For example, small geographical range, low population density and occupation of a high trophic level often figure as significant correlates of a high risk of extinction. Small changes in the factors that normally regulate population size can tip the balance towards extinction for species that are already less abundant or have restricted distributions. Many of these factors involve interactions with other species. Although the causes of current extinctions often lie outside the natural processes envisaged

by Darwin, many of them, such as the impact of invasive species, are enhanced versions of natural processes.

Recent work on West Indian birds provides an unusual opportunity to observe progressive stages of ecological and geographical contraction leading to extinction, using inferences from current geographical ranges and DNA-based phylogenies. Birds that colonize islands embark on sequential phases of range expansion and contraction, referred to as a taxon cycle[48,49]. Recent colonists are genetically indistinguishable from their mainland source populations, occupy wide geographical distributions and live mainly in lowland habitats[50]. Older, genetically differentiated, populations occur on one or a few islands, often restricted to forested environments at higher elevations. Each species provides a snapshot of this process, leading from the initial occupation of open, lowland habitats, to expansion into forested montane environments and exclusion from lowland habitats by new colonists. As they adapt to island interiors, older taxa also become isolated into small populations that are susceptible to local extinction. This process of cycling, with a consequent increase in the probability of extinction, fits well with Darwin's concept of extinction being driven by biotic interactions. Such support from the study of contemporary extinctions is understandably limited, but evidence should begin to accumulate more rapidly with the recent availability of molecular tools for investigating population history.

Darwin's proposal for the cause of extinction has yet to be fairly evaluated. He suggested that many taxa are driven to extinction by competition from ecologically similar but adaptively superior groups undergoing diversification. This core assumption of Darwin's explanation for macroevolution has little empirical support, mainly because the search for appropriate evidence has fallen through the gap between evolution and ecology; of course, pertinent evidence would strongly resist discovery under any circumstance.

Bringing divergence and extinction together

Central to Darwin's explanation for macroevolution is that the success of one group is gained at the expense of another. Palaeontological studies often reveal replacements in the fossil record, but their temporal, spatial and taxonomic resolution is generally limited. Research on faunal replacements tends to focus on biotic changes associated with major changes in the Earth's environment, such as the mid-Tertiary temperature decrease and the increasing aridity at temperate latitudes. For example, Christine Janis and co-workers documented the Miocene replacement of browsing mammals by grazers in North America and attributed it to replacement of forest by grasslands[51]. Jin Meng and Malcolm McKenna documented a similar Eocene–Oligocene replacement on the Mongolian plateau — where a perissodactyl-dominated mammalian fauna occupying a forested landscape was replaced by a rodent- and lagomorph-dominated grassland fauna — in association with the uplift of the Himalayas and the Tibetan plateau[52].

More direct evidence of continual background replacement of species comes from long fossil sequences with reasonable taxonomic resolution, such as the 40-million-year Palaeocene-to-Miocene record of pollen morphotypes from northwestern South America documented by Carlos Jaramillo and colleagues[53]. Although environment and overall diversity both varied over this period, the morphospecies composition of the flora turned over continually, even during periods of relative climate stability. More marked examples of such changes in the fossil record include the low-latitude replacement of gymnosperms by angiosperms during the mid-Cretaceous[54,55], and the post-Eocene replacement of non-passerine birds by passerines in Europe[56]. In both cases, older groups were replaced almost completely by more modern groups, leaving only sporadic relicts of the biotas — the Podocarpaceae and *Gnetum* in the case of gymnosperms, and swifts and woodpeckers, among others, in the case of European birds.

Additional indirect evidence for Darwin's theory of diversification and extinction comes from the observation that, following recovery from mass extinction events, species richness remains relatively stable. This pattern appears in the fossil record[57] and is evident in the absence of a correlation between clade age and contemporary diversity in several taxa[27,58]. This pattern suggests an underlying equilibrium between speciation and extinction during these intervals[24,29]. Perhaps speciation and extinction are random, unconnected events that balance out over time. However, models of random speciation and extinction processes suggest that the average time required for the complete replacement of species is approximately the product of the number of species and the average duration of individual species. The duration of species can be estimated directly from the fossil record[31] and, more recently, from the analysis of phylogenies[24,26]. These estimates, which typically fall between 1 million and 10 million years, are far too long for random speciation and extinction to account for observed species turnover rates[27]. Instead, more rapid turnover of species can be reconciled only when some lineages exhibit an excess of speciation and others an excess of extinction. Only a direct link between the two is needed to support Darwin's mechanism of macroevolution.

Conclusions

Although Darwin might have erred in some of the details of his principle of divergence, particularly the generally agreed starting point of reproductively isolated species, his basic idea has merit. The fundamental truth of his principle of divergence has emerged in different facets of evolutionary ecology, a field in which the same principle, in the form of character displacement or some models of sympatric speciation, was discovered independently in different contexts over a century after the publication of *On the Origin of Species*. Darwin's linking of extinction to diversification did not re-emerge as the study of extinction rose to prominence in conservation biology. Competitive replacement leading to extinction was once generally and uncritically accepted by palaeontologists before fading into the background after the discovery of mass extinctions. There is compelling evidence, however, of a role for biotic interactions in at least some extinction events, and a complementary relationship between divergence and extinction finds enough support for Darwin's proposal to merit further consideration as a viable link between microevolution and macroevolution.

Darwin's proposal carries a more general message for contemporary discussions of macroevolution, namely that microevolution alone cannot explain macroevolution. Understanding macroevolution requires the integration of ecology, evolution and the role of history in shaping the diversification or decline of lineages. Other investigators, most recently David Jablonski[59], have conveyed similar messages. Jablonski's vision is more complex than Darwin's and reflects the growth of ecology, evolution and palaeontology as disciplines since 1859, but it retains Darwin's emphasis on the presence of a biological filter that lies between microevolution and macroevolution and shapes the long-term consequences of evolutionary change. Jablonski concludes by calling for increased integration between fields to build a bridge between microevolution and macroevolution, and we concur with him. Mass extinctions and the large-scale expansions and contractions of clades in the fossil record are captivating but are only part of the story. Background extinctions are more elusive, but they must be considered in order to understand Darwin's mechanism of the turnover of species resulting from their evolution and interactions. Studies of extant populations — including mechanisms of population regulation, the contemporary causes of extinction and the causes of adaptive radiations — can yield important clues to factors that shape the history of life. Finally, information about the historical patterns of diversification of lineages can now be mined from molecular phylogenies, shedding light on the underlying causes of these patterns. It is the integration of information from the fossil record, the population and evolutionary dynamics of extant organisms, and phylogenetics that will provide the ultimate test of Darwin's bridge between microevolution and macroevolution.

Many people see *On the Origin of Species* as a beautiful fossil, but we view it as a living document that continues to offer insights and to enlighten modern research. It contains a wealth of ideas that have slipped through the cracks of the modern synthesis and, when appropriately updated, can provide inspiration for addressing some of the major unanswered questions in evolutionary biology. ■

1. Darwin, C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (John Murray, 1859).
This book is essential reading for those who wish not only to understand evolution in general but also to see the wealth of Darwin's original ideas that have yet to be tested.
2. Wallace, A. R. On the tendency of varieties to depart indefinitely from the original type. *J. Proc. Linn. Soc. (Zool.)* **3,** 53–62 (1858).
3. Rudwick, M. J. S. *Bursting the Limits of Time* (Univ. Chicago Press, 2005).
This book provides the best explanation for how it became evident that fossils represent the history of life and how the geological timescale was developed.
4. Mayr, E. Reasons for the failure of theories. *Phil. Sci.* **61,** 529–533 (1994).
5. Mayr, E. *Systematics and the Origin of Species* (Columbia Univ. Press, 1942).
6. Dobzhansky, T. *Genetics and the Origin of Species* (Columbia Univ. Press, 1937).
7. MacArthur, R. H. *Geographical Ecology* (Princeton Univ. Press, 1972).
8. MacArthur, R. H. & Levins, R. The limiting similarity, convergence, and divergence of coexisting species. *Am. Nat.* **101,** 377–386 (1967).
9. Vandermeer, J. H. Niche theory. *Annu. Rev. Ecol. Syst.* **3,** 107–132 (1972).
10. Chase, J. M. & Leibold, M. A. *Ecological Niches. Linking Classical and Contemporary Approaches* (Univ. Chicago Press, 2003).
11. Brown, W. L. & Wilson, E. O. Character displacement. *Syst. Zool.* **5,** 49–65 (1956).
This paper is a benchmark in the birth of evolutionary ecology and the first formal restatement of Darwin's principle of divergence.
12. Lack, D. *Darwin's Finches* (Cambridge Univ. Press, 1947).
13. Grant, P. R. Convergent and divergent character displacement. *Biol. J. Linn. Soc.* **4,** 39–68 (1972).
14. Losos, J. B. A phylogenetic analysis of character displacement in the Caribbean *Anolis* lizards. *Evolution* **44,** 558–569 (1990).
15. Schluter, D. & McPhail, J. D. Ecological character displacement and speciation in sticklebacks. *Am. Nat.* **140,** 85–108 (1992).
16. Losos, J. B. Ecological character displacement and the study of adaptation. *Proc. Natl Acad. Sci. USA* **97,** 5693–5695 (2000).
17. Simpson, G. G. *The Major Features of Evolution* (Columbia Univ. Press, 1953).
18. Schluter, D. *The Ecology of Adaptive Radiation* (Oxford Univ. Press, 2000).
In this paper, Schluter updates Simpson's concept of adaptive radiation and integrates it with the modern evidence for such radiations.
19. Harvey, P. H., May, R. M. & Nee, S. Phylogenies without fossils. *Evolution* **48,** 523–529 (1994).
20. Raup, D. M., Gould, S. J., Schopf, T. J. M. & Simberloff, D. S. Stochastic models of phylogeny and evolution of diversity. *J. Geol.* **81,** 525–542 (1973).
21. Nee, S., Mooers, A. O. & Harvey, P. H. Tempo and mode of evolution revealed from molecular phylogenies. *Proc. Natl Acad. Sci. USA* **89,** 8322–8326 (1992).
22. Nee, S., May, R. M. & Harvey, P. H. The reconstructed evolutionary process. *Phil. Trans. R. Soc. Lond. B* **344,** 305–311 (1994).
23. Pybus, O. G. & Harvey, P. H. Testing macro-evolutionary models using incomplete molecular phylogenies. *Proc. R. Soc. Lond. B* **267,** 2267–2272 (2000).
24. Nee, S. Birth–death models in macroevolution. *Annu. Rev. Ecol. Syst.* **37,** 1–17 (2006).
25. Harvey, P. H., Holmes, E. C. & Nee, S. Model phylogenies to explain the real world. *Bioessays* **16,** 767–770 (1994).
26. Ricklefs, R. E. Estimating diversification rates from phylogenetic information. *Trends Ecol. Evol.* **22,** 601–610 (2007).
27. Ricklefs, R. E. in *Speciation and Patterns of Diversity* (eds Butlin, R., Bridle, J. & Schluter, D.) 257–277 (Cambridge Univ. Press, 2009).
28. Ricklefs, R. E. Evolutionary diversification and the origin of the diversity–environment relationship. *Ecology* **87,** S3–S13 (2006).
29. Phillimore, A. B. & Price, T. D. Density-dependent cladogenesis in birds. *PLoS Biol.* **6,** 483–489 (2008).
30. Rabosky, D. L. & Lovette, I. J. Density-dependent diversification in North American wood warblers. *Proc. R. Soc. Lond. B* **275,** 2363–2371 (2008).
31. Stanley, S. M. *Macroevolution, Pattern and Process* (Freeman, 1979).
32. Wiens, J. J. & Donoghue, M. J. Historical biogeography, ecology, and species richness. *Trends Ecol. Evol.* **19,** 639–644 (2004).
33. Ricklefs, R. E., Schwarzbach, A. E. & Renner, S. S. Rate of lineage origin explains the diversity anomaly in the world's mangrove vegetation. *Am. Nat.* **168,** 805–810 (2006).
34. Latham, R. E. & Ricklefs, R. E. Global patterns of tree species richness in moist forests: energy-diversity theory does not account for variation in species richness. *Oikos* **67,** 325–333 (1993).
35. Judd, W. S., Sanders, R. W. & Donoghue, M. J. Angiosperm family pairs: preliminary phylogenetic analyses. *Harv. Pap. Bot.* **5,** 1–51 (1994).
36. Ricklefs, R. E. in *Tropical Rainforests: Past, Present, and Future* (eds Bermingham, E., Dick, C. W. & Moritz, C.) 16–40 (Univ. Chicago Press, 2005).
37. Rabosky, D. L., Donnellan, S. C., Talaba, A. L. & Lovette, I. J. Exceptional among-lineage variation in diversification rates during the radiation of Australia's most diverse vertebrate clade. *Proc. R. Soc. B* **274,** 2915–2923 (2007).
38. Barraclough, T. G., Vogler, A. P. & Harvey, P. H. Revealing the factors that promote speciation. *Phil. Trans. R. Soc. Lond. B* **353,** 241–249 (1998).
39. Freeman, P. W. Macroevolution in Microchiroptera: recoupling morphology and ecology with phylogeny. *Evol. Ecol. Res.* **2,** 317–335 (2000).
40. Ricklefs, R. E. Global diversification rates of passerine birds. *Proc. R. Soc. Lond. B* **270,** 2285–2291 (2003).
41. Ricklefs, R. E. Small clades at the periphery of passerine morphological space. *Am. Nat.* **165,** 651–659 (2005).
42. Raup, D. M. & Sepkoski, J. J. Mass extinctions in the marine fossil record. *Science* **215,** 1501–1503 (1982).
43. Raup, D. M. A kill curve for Phanerozoic marine species. *Paleobiology* **17,** 37–48 (1991).
Reference 42 presents the discovery of the most dramatic mass extinctions, whereas reference 43 presents a more general analysis of the distribution of extinction events throughout the fossil record.
44. Smith, J. T. & Roy, K. Selectivity during background extinction: Plio-Pleistocene scallops in California. *Paleobiology* **32,** 408–416 (2006).
45. Owens, I. P. F. & Bennett, P. M. Ecological basis of extinction risk in birds: habitat loss versus human persecution and introduced predators. *Proc. Natl Acad. Sci. USA* **97,** 12144–12148 (2000).
46. Jones, K. E., Purvis, A. & Gittleman, J. L. Biological correlates of extinction risk in bats. *Am. Nat.* **161,** 601–614 (2003).
47. Cardillo, M. *et al.* The predictability of extinction: biological and external correlates of decline in mammals. *Proc. R. Soc. Lond. B* **275,** 1441–1448 (2008).
48. Wilson, E. O. The nature of the taxon cycle in the Melanesian ant fauna. *Am. Nat.* **95,** 169–193 (1961).
49. Ricklefs, R. E. & Cox, G. W. Taxon cycles in the West Indian avifauna. *Am. Nat.* **106,** 195–219 (1972).
50. Ricklefs, R. E. & Bermingham, E. The concept of the taxon cycle in biogeography. *Glob. Ecol. Biogeogr.* **11,** 353–361 (2002).
51. Janis, C. M., Damuth, J. & Theodor, J. M. Miocene ungulates and terrestrial primary productivity: where have all the browsers gone? *Proc. Natl Acad. Sci. USA* **97,** 7899–7904 (2000).
52. Meng, J. & McKenna, M. C. Faunal turnovers of Palaeogene mammals from the Mongolian plateau. *Nature* **394,** 364–367 (1998).
53. Jaramillo, C., Rueda, M. J. & Mora, G. Cenozoic plant diversity in the Neotropics. *Science* **311,** 1893–1896 (2006).
54. Crane, P. R. & Lidgard, S. Angiosperm diversification and paleolatitudinal gradients in Cretaceous floristic diversity. *Science* **246,** 675–678 (1989).
55. Lidgard, S. & Crane, P. R. Angiosperm diversification and Cretaceous floristic trends; a comparison of palynofloras and leaf macrofloras. *Paleobiology* **16,** 77–93 (1990).
56. Mayr, G. The Paleogene fossil record of birds in Europe. *Biol. Rev.* **80,** 515–542 (2007).
57. Alroy, J. *et al.* Phanerozoic trends in the global diversity of marine invertebrates. *Science* **321,** 97–100 (2008).
58. Magallón, S. & Sanderson, M. J. Absolute diversification rates in angiosperm clades. *Evolution* **55,** 1762–1780 (2001).
59. Jablonski, D. Biotic interactions and macroevolution: extensions and mismatches across scales and levels. *Evolution* **62,** 715–739 (2008).
60. Rabosky, D. L. & Lovette, I. J. Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? *Evolution* **62,** 1866–1875 (2008).
61. Mayr, E. *The Growth of Biological Thought* (Harvard Univ. Press, 1982).
62. Gould, S. J. & Eldredge, N. Punctuated equilibria: the tempo and mode of evolution reconsidered. *Paleobiology* **3,** 115–151 (1977).
63. Gould, S. J. *The Structure of Evolutionary Theory* (Harvard Univ. Press, 2002).
64. Simpson, G. G. *Tempo and Mode in Evolution* (Columbia Univ. Press, 1944).
65. Wright, S. Character, change, speciation and higher taxa. *Evolution* **36,** 427–443 (1982).
This paper is ideal reading for those who wish to learn more about the debate on the relationship between microevolution and macroevolution.
66. Gould, S. J. *Ontongeny and Phylogeny* (Harvard Univ. Press, 1977).

Acknowledgements We thank M. Clark, N. Hughes, D. Rabosky and J. Sachs for comments on the manuscript. D.N.R. is supported by grants from the National Science Foundation (grant numbers DEB-0416085 and DEB-0623632).

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence should be addressed to D.N.R. (gupy@ucr.edu).

The nature of selection during plant domestication

Michael D. Purugganan[1] & Dorian Q. Fuller[2]

Plant domestication is an outstanding example of plant–animal co-evolution and is a far richer model for studying evolution than is generally appreciated. There have been numerous studies to identify genes associated with domestication, and archaeological work has provided a clear understanding of the dynamics of human cultivation practices during the Neolithic period. Together, these have provided a better understanding of the selective pressures that accompany crop domestication, and they demonstrate that a synthesis from the twin vantage points of genetics and archaeology can expand our understanding of the nature of evolutionary selection that accompanies domestication.

Domestication is a complex evolutionary process in which human use of plant and animal species leads to morphological and physiological changes that distinguish domesticated taxa from their wild ancestors[1]. It is one of the most important technological innovations in human history and was the linchpin of the Neolithic revolution 13,000–10,000 years ago, in which groups of hunter-gatherers formed the sedentary agricultural societies that ultimately gave rise to current human cultures[2]. Domestication gave rise to food surpluses, and this led to craft specializations, art, social hierarchies, writing, urbanization and the origin of the state[2].

As a process of recent, rapid species evolution, domestication was of great interest to Charles Darwin when he formulated his thesis on the origin of species through natural selection[3,4]. Evolutionary biologists, however, tend to view domestication as a special class of species diversification, distinct from species divergence through natural selection in the wild[5]. Yet domestication can also be seen as a type of plant–animal co-evolution, conceptually similar to examples of evolutionary diversification driven by other multispecies interactions[6,7]. Indeed, the spread of crop species, which today dominate landscapes across the planet, attests to the increased fitness of domesticated plant taxa and suggests that domestication is one of the most successful of all plant–animal mutualisms. Moreover, fungal species have been domesticated by ants[8] and beetles[9], so domestication is not specific to *Homo sapiens*. Nevertheless, the role of human culture, including the intentional manipulation of plants as sources of delayed food returns, drives the domestication process in distinctive ways.

The use of domestication as a model for the evolutionary process stems from an understanding of events associated with the origins of crop species (starting some 13,000 years ago) and from precise knowledge of the selective pressures experienced by domesticated taxa, which can be gleaned from archaeological data and ethnographic studies of traditional farming societies and hunter-gatherers. Archaeology, in particular, can establish a fossil framework in which changes in phenotypes can be tracked in space and time and dated relatively precisely, allowing the microevolutionary dynamics that accompany species diversification to be traced. Genetic information on crop species also provides a molecular framework in the study of this co-evolutionary process[10,11], linking selective mechanisms inferred from archaeological studies to the genes that drove the origin and diversification of crop plant species.

In this Review, we discuss recent archaeological work that reveals the mechanisms of the adaptation of crop plants to cultivation in agricultural environments and human cultures, and we describe genetic and genomic studies into the nature of adaptive selection in the genomes of crop species. The focus on both genetic and archaeological insights provides a clear picture of the selective pressures that accompany crop origins and diversification. The view from these two vantage points can increase understanding of the nature of the evolutionary selection that accompanies plant domestication.

Cultivation and the rise of domesticated species

Humans were initially foragers and for a long time ate wild cereals, as well as seeds and nuts. Evidence from the Ohalo II archaeological site in Israel, for example, shows that wild wheat and barley were used at least 10,000 years before the advent of cultivation[12]. Beginning in the Epipalaeolithic and into the Neolithic period (13,000 to 11,000 years ago), however, foraging gave way to cultivation, representing a shift of labour investment to just a few plant species as food sources[13,14] and a change in human behavioural ecology that selected for recurrent adaptations, leading to domestication[13,15,16].

Archaeological evidence suggests that hunter-gatherer groups independently began cultivating food plants in 24 regions, and grain crops (mostly grasses) were the focus of early cultivation in perhaps 13 regions (Fig. 1). Several traits evolved that result in the distinctive morphologies and physiologies that distinguish many domesticated plant species from their wild ancestors. Domestication traits differ between plants, depending on the way they are used, for example whether they are grown for fruits and vegetative organs or edible seeds[15,16]. Changes in grain crops — including various seed traits, a shift to an annual life cycle, increased selfing rates and reduced lateral branching or tillering — are well documented and are the focus of this review.

Adaptations in cereal crops that evolve under human cultivation fall into two main types: responses that lead to successful germination with increased soil disturbance and depth of burial, and those that make harvesting easier[17,18]. Several of these adaptations may proceed from unconscious selective pressures that act in the same way as natural selection in the wild, albeit under the imposed agroecological environments of the cultivated field[13,17,18].

[1]Department of Biology and Centre for Genomics and Systems Biology, 100 Washington Square East, New York University, New York 10003, USA. [2]Institute of Archaeology, University College London, 31–34 Gordon Square, London WC1H 0PY, UK.

Two such traits are increases in grain size and loss of seed shattering, both of which are well represented in archaeological data (Figs 2 and 3). An increase in seed size is an early adaptive response to human cultivation, and greater seed size is strongly correlated with larger seedlings in many cereal and legume species[19]. Comparative ecological studies show that larger seeds have advantages under certain kinds of competition, including deeper burial in soil[19,20]. Larger seeds are therefore likely to be selected for open environments, where bigger seedlings are advantageous given the deeper burial in soils disturbed by human tillage[13,17].

Measurable increases in grain size among barley and einkorn wheat in archaeological samples have been observed in the Jerf el Ahmar site in the Upper Euphrates in the Early Neolithic, some 500–1,000 years after the beginnings of cultivation[13]. This trait also shows relatively rapid local evolution (less than 500 years) in rye, as shown by the presence of enlarged grains at Abu Hureyra in Syria in the Late Pleistocene[21,22]. By contrast, archaeological evidence for rice suggests that changes in grain size[23] and husk phytoliths[24] were most intensive over a longer period, between 9,000 and 5,500 years ago. Archaeological data also suggest a period of grain size enlargement of one to two millennia[25] in a minor domesticated millet in Japan, *Echinochloa crusgalli* var. *utilis*.

The reduction of natural seed dispersal, which includes the loss of seed shattering, is the other key trait that is selected under cultivation[13,16,17]. Non-shattering is often regarded as the hallmark of domestication in most seed crops because it renders a plant species primarily dependent on humans for survival and propagation. The loss of shattering in cereal crop species generally arises from the absence of an abscission layer at the spikelet base[26] (Fig. 3) and has been attributed to the early use of sickle tools that select for the retention of seed on the plant after harvesting[27,28]. Three genes favouring seeds that do not shatter have been isolated: the rice gene *sh4* (ref. 26), which is similar to the genes encoding MYB-like transcription factors in maize; the rice quantitative trait locus (QTL) *qSH1*, which encodes a homeobox-containing protein[29]; and the wheat gene *Q*, which is similar to genes of the *AP2* family in other plants[30].

The archaeological record provides evidence that seed size enlargement can occur before the loss of shattering. In wheat and barley, archaeological studies demonstrate that an increase in grain size was followed by the fixation of non-shattering rachises[13]. There is also evidence from China that increased grain size was being selected under cultivation about 8,000 years ago[23,31], before there was a general increase in the frequency of spikelet types associated with loss of shattering. These findings indicate that seed size enlargement arises during the early history of cultivation but before the completion of plant domestication, and that an increase in grain size may be one of the first traits to experience selective pressures associated with human cultivation.

This evolutionary sequence of seed size increase before the rise of non-shattering is by no means universal, however. Grain size enlargement in pearl millet, for example, did not occur until 2,000 years after domestication[12], although evidence from Libya and India indicates that selection for grain enlargement in pearl millet occurred several times[13]. Data from India and sub-Saharan Africa also suggest that the increase in seed size in pulses was delayed by millennia after the beginning of cultivation and may not have been part of the initial domestication[32,33].

The evolution of non-shattering was also a slow process, despite modelling studies demonstrating that it can evolve in less than 100 years with the use of sickle tools[28,34]. Archaeobotanical remains of wheat and barley ears indicate that the fixation of non-shattering rachises took about 2,000 years[13,35] (Fig. 4). Recent efforts to recover rice spikelet bases, which preserve evidence for the non-shattering, domesticated, type, also show that fixation of this trait was surprisingly late in China. As recent as 7,500 to 6,500 years ago, only around half of the rice from the Lower Yangtze had the domesticated, non-shattering, morphology, and

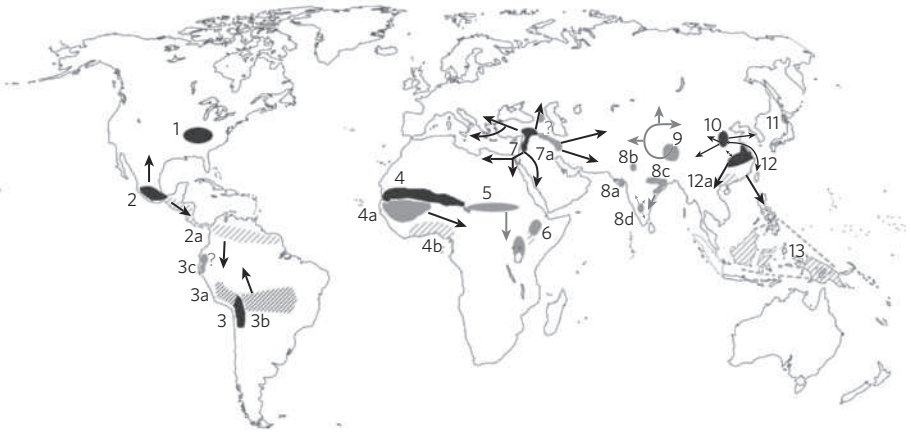


Figure 1 | Centres of plant domestication. Solid-shaded areas and hatched areas indicate regions of important seed-crop domestication and vegecultural crops, respectively. Accepted primary domestication centres are shown in black, and potentially important secondary domestication centres are shown in grey. Arrows indicate major trajectories of spread of agriculture and crops out of some centres. Areas are numbered, and examples of crop species and the year by which they were domesticated in each area are as follows: 1, eastern North America (*Chenopodium berlandieri*, *Iva annua* and *Helianthus annuus*, 4,500–4,000 years before present (yr BP)); 2, Mesoamerica (*Cucurbita pepo*, 10,000 yr BP; *Zea mays*, 9,000–7,000 yr BP); 2a, northern lowland neotropics (*Cucurbita moschata*, *Ipomoea batatas*, *Phaseolus vulgaris*, tree crops, 9,000–8,000 yr BP); 3, central mid-altitude Andes (*Chenopodium quinoa*, *Amaranthus caudatus*, 5,000 yr BP); 3a, north and central Andes, mid-altitude and high altitude (*Solanum tuberosum*, *Oxalis tuberosa*, *Chenopodium pallidicaule*, 8,000 yr BP); 3b, lowland southern Amazonia (*Manihot esculenta*, *Arachis hypogaea*, 8,000 yr BP); 3c, Ecuador and northwest Peru (*Phaseolus lunatus*, *Canavalia plagiosperma*, *Cucurbita ecuadorensis*, 10,000 yr BP; the question mark indicates that there is some question of the independence of crop origins of this centre from 3, 3a and 3b); 4, West African sub-Sahara (*Pennisetum glaucum*, 4,500 yr BP); 4a, West African savanna and woodlands (*Vigna unguiculata*, 3,700 yr BP; *Digitaria exilis*, *Oryza glaberrima*, <3,000 yr BP); 4b, West African rainforests (*Dioscorea rotundata*, *Elaeis guineensis*, poorly documented); 5, east Sudanic Africa (*Sorghum bicolor*, >4,000 yr BP?); 6, East African uplands (*Eragrostis tef*, *Eleusine coracana*, 4,000 yr BP?) and lowland vegeculture (*Dioscorea cayenensis*, *Ensete ventricosum*, poorly documented); 7, Near East (*Hordeum vulgare*, *Triticum* spp., *Lens culinaris*, *Pisum sativum*, *Cicer arietinum*, *Vicia faba*, 13,000–10,000 yr BP); 7a, eastern fertile crescent (additional *Hordeum vulgare* and, 9,000 yr BP, also goats); 8a, Gujarat, India (*Panicum sumatrense*, *Vigna mungo*, 5,000 yr BP?); 8b, Upper Indus (*Panicum sumatrense*, *Vigna radiata*, *Vigna aconitifolia*, 5,000 yr BP); 8c, Ganges (*Oryza sativa* subsp. *indica*, 8,500–4,500 yr BP); 8d, southern India (*Brachiaria ramosa*, *Vigna radiata*, *Macrotyloma uniflorum*, 5,000–4,000 yr BP); 9, eastern Himalayas and Yunnan uplands (*Fagopyrum esculentum*, 5,000 yr BP?); 10, northern China (*Setaria italica*, *Panicum miliaceum*, 8,000 yr BP; *Glycine max*, 4,500 yr BP?); 11, southern Hokkaido, Japan (*Echinochloa crusgalli*, 4,500 yr BP); 12, Yangtze, China (*Oryza sativa* subsp. *japonica*, 9,000–6,000 yr BP); 12a, southern China (*Colocasia*, *Coix lachryma-jobi*, poorly documented, 4,500 yr BP?); 13, New Guinea and Wallacea (*Colocasia esculenta*, *Dioscorea esculenta*, *Musa acuminata*, 7,000 yr BP).

there is evidence of only a gradual rise in frequency (ref. 36, and D.Q.F., L. Qin, Y. Zheng, Z. Zhao, X. Chen, L. Hosoya and G. Sun, unpublished observations) (Fig. 4).

The rate of increase of the non-shattering, domesticated, form of barley, einkorn wheat and rice can be estimated from the archaeological record and reflects the strength of selection on this trait (Fig. 4). Remarkably, these different crop species have similar rates of phenotypic selection, with non-shattering forms increasing at a rate of 0.03–0.04% per year. This suggests that selection on the loss of seed shattering in cereal crop species, and hence on domesticated morphologies, is similar across different taxa, geographical origins and time periods. The slow rate of evolution of this trait also implies weak selective pressures[13], especially compared with the relatively rapid rates of grain size evolution. The slow rise in non-shattering has been attributed to continued gene flow into proto-domesticates as humans continued to gather wild grain[37], leading to a period of metastable equilibrium in which adaptive diversity for harvesting and natural shattering coexisted in early crop populations.

From these and other archaeological data, it is clear that the origins of domesticated plant species were not single events but an extended multistage process in which traits arose sequentially over several thousand years to create the phenotypic assemblage that characterizes domesticated species today. The archaeological results are in contrast to molecular evolutionary studies, which have invariably assumed rapid, single origins for domestic species[38–40], when in reality these taxa may have evolved in a stepwise manner over several millennia.

Cultural and ecological adaptation

Selection after domestication has led to the immense diversity in varieties that characterizes many domesticated plant species, which, as Darwin pointed out, can exceed the range of phenotypic variation in their wild ancestors[6]. Selection for crop diversification leads to local adaptation, driven by human groups developing varieties with preferred cooking or processing qualities, the ability to grow in new environments, or desirable visual or gustatory features. Unlike domestication traits, however, selection for crop diversification may involve a greater level of conscious selection, as human cultures deliberately chose varieties with the desired characteristics.

One result of such cultural selective pressure is the development of crops that are less labour intensive to process after harvesting. In some species, this selective pressure was not local but species-wide; evolution in maize, for example, turned the hard podcase of teosinte into easily milled, but still glumed, pod-corns and then into naked-grained cobs[41,42]. This trait is controlled in part by the domestication gene *teosinte glume architecture* (*tga1*), which encodes a transcriptional activator belonging to the SBP family[42,43]. In other species, however, such as wheat and barley, local selection has resulted in diverse varieties. Recurrent evolution resulted in free-threshing sorghum races in northeast, west and southern Africa and India[44,45], and molecular data indicate that free-threshing tetraploid and hexaploid wheats evolved from ancestral glume wheats at least twice[46,47]. Archaeobotanical evidence also reveals

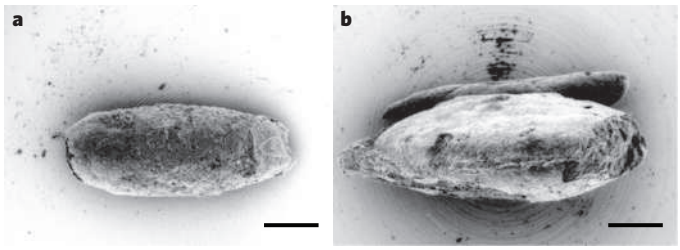


Figure 2 | Evolution of grain-size increases in the archaeological record. Scanning electron micrographs of the dorsal view of wild (**a**) and domesticated (**b**) einkorn wheat from Abu Hureyra. The wild einkorn wheat is from the pre-pottery Neolithic B (~9,300 yr BP), whereas the domesticated einkorn is from the latest levels of the site, some 8,000 yr BP or later. Scale bars, 1 mm.

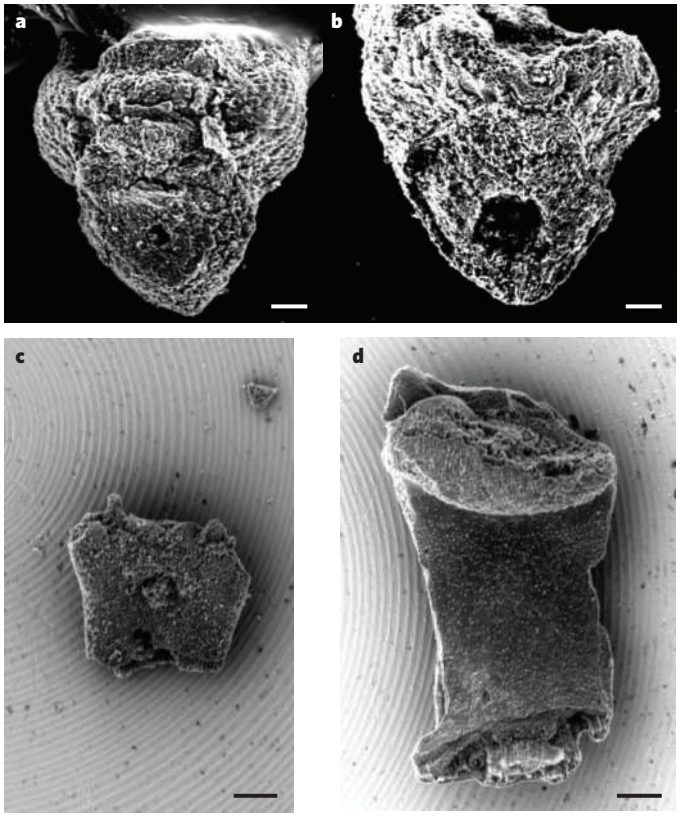


Figure 3 | The evolution of non-shattering seeds in the archaeological record. a, **b**, Scanning electron micrographs of charred wild (**a**) and domesticated (**b**) rice spikelet bases from the Tian Luo Shan site in China, ~6,700 yr BP. Scale bars, 100 μm. **c**, Wild-type barley rachis with abscission scar from Dhra, Jordan, ~11,000 yr BP. Scale bar, 250 μm. **d**, Typical carbonized rachis of domesticated barley with protruding rachis attachment at the top, from Kawa, Sudan, ~2,500 yr BP. Scale bar, 250 μm. (Panel **c** courtesy of S. Colledge, University College London, UK.)

that both tetraploid and hexaploid free-threshing wheats were present, although rare, during the later pre-pottery Neolithic in Syria and Turkey by 9,500 years ago[18,48]. Despite selective pressure for free-threshing varieties, some cultures still have a preference for hulled varieties, perhaps for improved storability or more reliable germination. As a result, hexaploid glume wheats evolved from free-threshing hexaploid wheats several times in various regions of western Eurasia[18,49].

Similarly, naked barley evolved a reduction in the tight-fitting hull and was also present in the pre-pottery Neolithic Near East[18,48]. Prehistoric Europe also had regionally varied patterns of reliance on naked barleys with six-rowed spikelets of grain[48]. Nevertheless, some cultures and periods have shown a clear preference for hulled barley, including 3,000 years ago during the Late Bronze Age in parts of northern and western Europe[50], where it may be associated with new forms of cooking or beer brewing.

Cultural selection associated with distinct food preferences is also fairly common, and this includes maize and sorghum varieties that have evolved tough pericarps and are used for making popcorn[51] and the evolution of the distinct fragrance of basmati and jasmine rice[52]. In the latter case, an increase in the levels of 2-acetyl-1-pyrroline results in part from a non-functional allele of *sk2*, which encodes a betaine aldehyde dehydrogenase, producing the distinctive aroma prized by south Asian, Thai and Iranian cultures[52].

One of the best-studied examples of post-domestication selection involves the preference for sticky cooked cereals in east Asian cultures, a property that arises from a reduction in amylose levels in grain starch. Eight cereal species in China, Korea, Japan and northern Southeast Asia are known to have sticky glutinous varieties[53], and this cultural taste may have its origins in hunter-gatherers processing starchy nuts

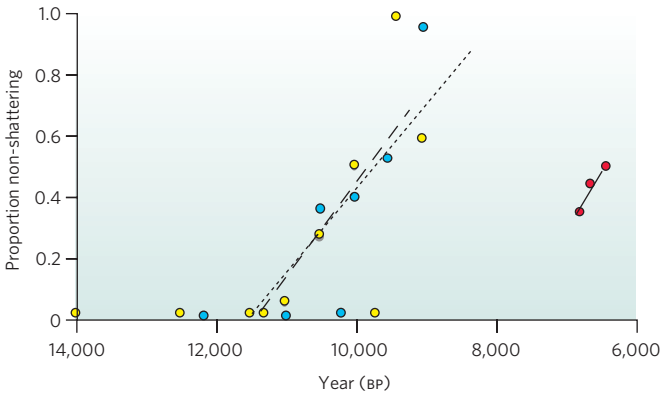


Figure 4 | Frequency of non-shattering, domesticated, forms of barley, wheat and rice in the archaeological record. The barley and wheat data are from Near Eastern sites[13], and the rice data are from the Lower Yangtze (D.Q.F., L. Qin, Y. Zheng, Z. Zhao, X. Chen, L. Hosoya and G. Sun, unpublished observations). The wheat data do not include the Kosak Shimali site[35], as it represents a later date and has a high proportion of indeterminate (damaged) rachises. The rate of increase in frequency is calculated from a linear model, with R^2 values of 0.48 for barley (yellow, long-dashed line), 0.66 for wheat (blue, short-dashed line) and 0.85 for rice (red, solid line).

and tubers[54]. The most widespread of these cereals is glutinous rice; this results from an intron 1 splice-donor mutation at the *waxy* gene, which encodes a starch-granule-bound starch synthase[55–57]. This *waxy* splice-donor mutation originated once in domesticated tropical *japonica* rice, probably in Southeast Asia[56]. By contrast, recent work on the *waxy* locus of foxtail millet has found three distinct mutations with differing geographical distributions that all produce sticky millets[58]. This parallel evolution reinforces the possibility of evolutionary genetic constraint and independent selection acting multiple times at similar loci during evolutionary diversification.

Crop species also underwent range expansions after domestication, through a combination of human migrations and the adoption of crops obtained by trade, which in many cases spread domesticated plants far from their centres of origin. Climatic and other ecological constraints control the rates and extents of domesticated migrations, as illustrated by the case of early Near Eastern crops such as wheat, barley, lentils, peas and chickpeas. These crops had spread rapidly east as far as Pakistan[59] and west through Greece by about 9,000 years ago, had reached the Balkans 8,000 years ago, and approximately 500 years later had spread to Italy, Spain and Portugal[48]. North of the Balkans and through the Carpathian Mountains, however, species spread was delayed by about 1,000 years, a pause attributed to the need to evolve either a vernalization response to cope with cold winters or photoperiod-neutral varieties that could be grown in summer[60,61]. The evolution of wheat and barley varieties that could expand to these northern latitudes allowed their spread, but species such as lentils and chickpeas did not adapt[62]. Similarly, widespread use of wheat and barley in mountainous Kashmir and northern Pakistan began about 5,000 years ago, three millennia after they arrived in the neighbouring Indus valley[59].

The nature of selection in crop genomes

Early QTL mapping studies suggest that many traits affected by domestication and diversification are controlled by just a few genes, some of which have large effects[63], although in sunflowers small-to-moderate effects are more common[64]. Comparative QTL mapping suggests that selection may sometimes have acted on the same loci several times, but in other cases different genes have been affected, for example in selection for non-shattering mutations in maize, rice and wheat[65]. Nevertheless, analysis of QTLs for disparate domestication traits reveals that selection need act on only a few genomic regions to achieve domestication, as indicated by studies in rice[66] and wheat[67]. Recent population-genomics studies, however, provide a contrasting view. The number of genes associated with crop domestication and diversification in maize, for example, is large, with 2–4% of genes in the genome showing evidence of selection[38]. Furthermore, a recent study of genome-wide nucleotide polymorphism in rice suggests that domestication may have affected the entire genome, possibly because selection acted on a large number of loci[39].

Recent successes in the isolation of genes underlying crop domestication and diversification have improved the ability to examine the genetic basis of the evolution of domesticated species. Molecular genetic studies have so far identified 9 domestication genes in plants, as well as 26 other loci known to underlie crop diversity associated with human cultural preferences or different agricultural environments[10,53,68,69]. Of the nine domestication loci, eight encode transcriptional activators[10], including the rice *shattering* genes *sh4* (ref. 26) and *qSH1* (ref. 29), maize *tb1* (which is involved in plant architecture)[70], and the *AP2*-like wheat gene *Q* (which is involved in inflorescence structure)[30]. Of the crop-diversification genes whose molecular functions are known, however, more than half encode enzymes[10,53,68,69]. Domestication, then, seems to be associated with changes in transcriptional regulatory networks, whereas crop diversification involves a larger proportion of enzyme-encoding loci. Moreover, both regulatory promoter changes and amino-acid changes (or disrupted coding sequences) are responsible for relevant evolved phenotypes[10], and loss-of-function alleles seem to be affected more by diversification than domestication[10].

Selective sweeps[71] — that is, the reductions of nucleotide variation that result from strong selective pressures acting on particular loci — have been observed in genes associated with domestication or diversification phenotypes. The best example of a domestication-related selective sweep occurs at the maize gene *tb1*, which is involved in the suppression of axillary-branch formation, where a selective sweep ~90 kb in length across the promoter and proximal intergenic region has been observed[72] (Fig. 5). Other examples include a 600-kb selective sweep[73] at the maize gene *Y1*, involved in the yellow-kernel phenotype, and a 260-kb sweep in the rice gene *waxy*, which is associated with low-amylose rice in north-east Asian cultivars[57] (Fig. 5).

The strength of selection on the loci affected by domestication and diversification can be estimated from the physical extent of a sweep[57]. Using this approach, inferred selection coefficients, *s*, for cases of natural selection range from 0.02 to 0.70, the latter being associated with drug resistance in the parasite *Plasmodium falciparum*[57]. For maize *tb1*, which is the only domestication locus for which the selection coefficient has been estimated, *s* is in the range 0.05–0.20, comparable to cases of natural selection. By contrast, the crop diversification genes *waxy* in rice and *Y1* in maize have higher selection coefficients ($s>1$). This may simply reflect differences in recombination rates or the relatively recent selection for diversification compared with domestication, but it may also suggest that stronger selection accompanies post-domestication diversification of crop phenotypes[57].

Combining phylogeographical analysis with molecular genetic analysis allows researchers to determine which geographical regions gave rise to various traits and to chronicle the spatial spread of alleles subject to Darwinian selection. The analysis at the rice *waxy* gene, for example, demonstrates that the causal glutinous rice polymorphism had a localized origin in Southeast Asia and subsequently spread to northeastern Asia, where it is culturally valued[56]. Another study reveals that the loss-of-function allele of the rice *Rc* gene, which encodes a basic helix–loop–helix protein and gives rice the white pericarp favoured by most Asian cultures, probably originated in tropical *japonica* rice in Southeast Asia but spread into the *indica* rice variety that is found primarily in southern Asia[74].

Solving an obscure problem

As Darwin noted in *On the Origin of Species*[3], adaptation to a cultivated environment or human cultural preferences — "man's use or fancy" — is one of the hallmarks of domesticated species, an observation that served him well when he formulated his theory of natural selection. In the 150 years since, both archaeology and genetics have provided crucial insights into the nature and timing of domestication, which

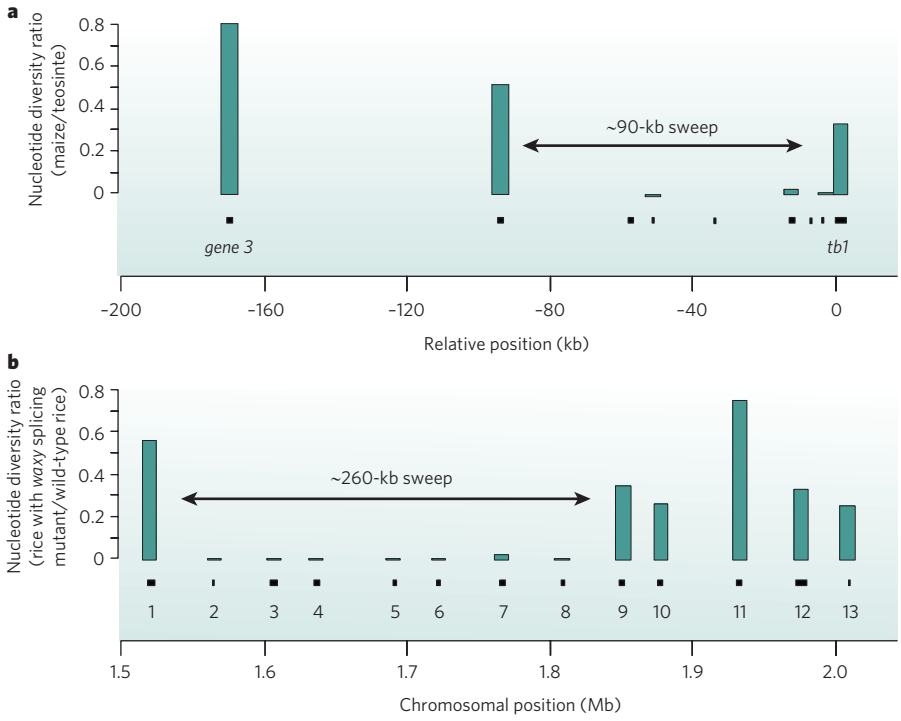


Figure 5 | Examples of selective sweeps at the maize gene *tb1* and at the rice locus *waxy*. Selection leads to a sweep region that appears as a region of reduced nucleotide diversity in the genome. Black boxes indicate the positions of sequenced regions relative to the maize gene *tb1* (**a**) and the physical positions of the sequenced regions on rice chromosome 6 (**b**). Green rectangles show the ratios of nucleotide diversity at these regions between maize and teosinte (**a**) and between rice accessions with the splicing mutant of *waxy* and wild-type rice (**b**). Sequenced regions in **b** include hypothetical genes (1, 2, 6 and 9), expressed sequence tags (4 and 10), the *waxy* locus (7), and loci that encode a receptor-like protein (3), glycosyl hydrolase (5), transketolase (8), enolase (11), a kinesin-like protein (12) and ribosome-binding factor A (13). The estimated extent of the selective sweep is indicated by the arrows.

underpinned the Neolithic revolution and gave rise to agrarian societies (which became the dominant *H. sapiens* cultures). The extent to which domestication differs from natural selection on wild species remains to be understood, as well as whether these processes are simply part of a continuum of selective regimes, differing merely in type and intensity but conceptually identical. To what extent does parallel evolution affect the same genes across multiple species during domestication and diversification? How can archaeobotanical evidence constrain molecular, evolutionary and genetic models of domestication by providing more realistic estimates of the time span for the evolution of these domestication traits, as well as the timing and strength of selection events? Can we combine archaeological data with genomic information to track the timing and geographical origins of traits? Answering these and other questions requires continued work to isolate more domestication and diversification genes, and more-systematic archaeological studies are needed to provide a comprehensive picture of the origins and dynamics of domestication.

Even now, 150 years after the publication of Darwin's most influential work, domesticated species are some of the best-studied examples of evolutionary diversification, and archaeological and genetic investigations of this unique case of plant–animal co-evolution have provided some of the most comprehensive views of species diversification. The case Darwin made in *On the Origin of Species* remains true: the study of domestication is central to understanding the nature of what Darwin referred to as "the obscure problem" of natural selection. ■

1. Hancock, J. F. Contributions of domesticated plant studies to our understanding of plant evolution. *Ann. Bot. (Lond.)* **96,** 953–963 (2005).
2. Diamond, J. Evolution, consequences and future of plant and animal domestication. *Nature* **418,** 700–707 (2002).
3. Darwin, C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (John Murray, 1859).
4. Darwin, C. *The Variation of Animals and Plants under Domestication* (John Murray, 1868).
5. Darwin, C. & Wallace, A. R. On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. *J. Proc. Linn. Soc. (Zool.)* **3,** 46–50 (1858).
6. Rindos, D. *The Origins of Agriculture: An Evolutionary Perspective* (Academic, 1984).
7. Zeder, M. A., Emshwiller, E., Smith, B. D. & Bradley, D. G. Documenting domestication: the intersection of genetics and archaeology. *Trends Genet.* **22,** 139–155 (2006).
8. Schultz, T. R. & Brady, S. G. Major evolutionary transitions in ant agriculture. *Proc. Natl Acad. Sci. USA* **105,** 5435–5440 (2008).
9. Farrell, B. D. *et al.* The evolution of agriculture in beetles (Curculionidae: Scolytinae and Platypodinae). *Evolution* **55,** 2011–2027 (2001).
10. Doebley, J. F., Gaut, B. S. & Smith, B. D. The molecular genetics of crop domestication. *Cell* **127,** 1309–1329 (2006).
11. Burke, J. M., Burger, J. C. & Chapman, M. A. Crop evolution: from genetics to genomics. *Curr. Opin. Genet. Dev.* **17,** 525–532 (2007).
12. Kislev, M. E., Nadel, D. & Carmi, I. Epipalaeolithic cereal and fruit diet at Ohalu II, Sea of Galilee, Israel. *Rev. Palaeobot. Palyn.* **73,** 161–166 (1992).
13. Fuller, D. Q. Contrasting patterns in crop domestication and domestication rates: recent archaeobotanical insights from the Old World. *Ann. Bot. (Lond.)* **100,** 903–924(2007).
This paper synthesizes and compares quantitative data for morphological change across time from archaeologically dated subfossil crop remains.
14. Harris, D. R. in *Foraging and Farming: The Evolution of Plant Exploitation* (eds Harris, D. R. & Hillman, G. C.) 11–26 (Routledge, 1989).
15. Hammer, K. Das Domestikationssyndrom. *Kulturpflanze* **32,** 11–34 (1984).
16. Smith, B. D. in *Documenting Domestication* (eds Zeder, M. A., Bradely, D. G., Emshwiller, E. & Smith, B. D.) 15–24 (Univ. California Press, 2006).
17. Harlan, J. R., De Wet, J. M. J. & Price, E. G. Comparative evolution of cereals. *Evolution* **27,** 311–325 (1973).
18. Zohary, D. & Hopf, M. *Domestication of Plants in the Old World* (Oxford Univ. Press, 2000).
19. Baskin, C. & Baskin, J. M. *Seeds: Ecology, Biogeography and Evolution of Dormancy and Germination* (Academic, 2001).
20. Westoby, M., Leishman, M. & Lord, J. Comparative ecology of seed size and dispersal. *Phil. Trans. R. Soc. Lond. B* **351,** 1309–1317 (1996).
21. Hillman, G. C. in *Village on the Euphrates: From Foraging to Farming at Abu Hureyra* (eds Moore, A. M. T., Hillman, G. C. & Legge, A. J.) 327–398 (Oxford Univ. Press, 2000).
22. Hillman, G. C., Hedges, R., Moore, A. M. T., Colledge, S. & Pettitt, P. New evidence of Late Glacial cereal cultivation at Abu Hureyra on the Euphrates. *Holocene* **11,** 383–393 (2001).
23. Fuller, D. Q. *et al.* Presumed domestication? Evidence for wild rice cultivation and domestication in the fifth millennium BC of the Lower Yangtze region. *Antiquity* **81,** 316–331 (2007).
24. Zhao, Z. The Middle Yangtze region in China is one place where rice was domesticated: phytolith evidence from the Diaotonghuan Cave, Northern Jaingxi. *Antiquity* **72,** 885–897 (1998).
25. Crawford, G. *Paleoethnobotany of the Kameda Peninsula Jomon* (Museum of Anthropology, Univ. Michigan, 1983).
26. Li, C. B., Zhou, A. L. & Sang, T. Rice domestication by reducing shattering. *Science* **311,** 1936–1939 (2006).
This paper reports the first molecular isolation of a loss-of-shattering gene in a cereal crop species, a trait that is the hallmark of domestication in seed crops.
27. Wilke, P. J., Bettinger, R., King, T. F. & O'Connell, J. F. Harvest selection and domestication in seed plants. *Antiquity* **46,** 203–209 (1972).
28. Hillman, G. & Davies, M. S. Domestication rates in wild wheats and barley under primitive cultivation. *Biol. J. Linn. Soc.* **39,** 39–78 (1990).
29. Konishi, S. *et al.* An SNP caused loss of seed shattering during rice domestication. *Science* **312,** 1392–1396 (2006).
30. Simons, K. J. *et al.* Molecular characterization of the major wheat domestication gene *Q*. *Genetics* **172,** 547–555 (2006).
31. Liu, L., Lee, G.-A., Jiang, L. & Zhang, J. Evidence for the early beginning (c. 9000 cal. BP) of rice domestication in China: a response. *Holocene* **17,** 1059–1068 (2007)
32. Fuller, D. Q. & Harvey, E. L. The archaeobotany of Indian pulses: identification, processing and evidence for cultivation. *Environ. Archaeol.* **11,** 219–246 (2006).
33. D'Andrea, A. C., Kahlheber, S., Logan, A. L. & Watson, D. J. Early domesticated cowpea (*Vigna unguiculata*) from Central Ghana. *Antiquity* **81,** 686–698 (2007).
34. D'Ennequin, M. L. T., Toupance, B., Robert, T., Godelle, B. & Gouton, P. H. Plant domestication: a model for studying the selection of linkage. *J. Evol. Biol.* **12,** 1138–1147 (1999).

35. Tanno, K. I. & Willcox, G. How fast was wild wheat domesticated? *Science* **311,** 1886 (2006).
36. Zheng, Y., Sun, G. & Chen, X. Characteristics of the short rachillae of rice from archaeological sites dating to 7000 years ago. *Chin. Sci. Bull.* **52,** 1654–1660 (2007).
37. Willcox, G., Fornite, S. & Herveux, L. Early Holocene cultivation before domestication in northern Syria. *Veg. Hist. Archaeobot.* **17,** 313–325 (2008).
This paper reports assemblages of morphologically wild cereals and early weed assemblages from two sites that indicate that wheat and barley were cultivated for more than 1,000 years before morphological domestication.
38. Wright, S. *et al*. The effects of artificial selection of the maize genome. *Science* **308,** 1310–1314 (2005).
This paper reports a genome-wide screen for genes that display a signature of positive selection associated with crop domestication and diversification.
39. Caicedo, A. L. *et al.* Genome-wide patterns of nucleotide polymorphism in domesticated rice. *PLoS Genet*. **3,** 1289–1299 (2007).
40. Eyre-Walker, A., Gaut, R. L., Hilton, H., Feldman, D. L. & Gaut, B. S. Investigation of the bottleneck leading to the domestication of maize. *Proc. Natl Acad. Sci. USA* **95,** 4441–4446 (1998).
41. Iltis, H. Homeotic sexual translocations and the origin of maize (*Zea mays*, Poaceae): a new look at an old problem. *Econ. Bot.* **54,** 7–42 (2000)
42. Wang, H. *et al.* The origin of the naked grains of maize. *Nature* **436,** 714–719 (2005).
43. Dorweiler, J., Stec, A., Kermicle, J. & Doeble, J. Teosinte-glume-architecture — a genetic locus controlling a key step in maize evolution. *Science* **262,** 233–235 (1993).
44. Harlan, J. R. & Stemler, A. B. in *The Origins of African Plant Domestication* (eds Harlan, J. R., De Wet, J. M. J. & Stemler, A. B.) 465–478 (Mounton, 1976).
45. DeWet, J. M. Systematics and evolution of sorghum. *Am. J. Bot.* **65,** 477–484 (1978).
46. Giles, R. & Brown, T. *GluDy* allele variations in *Aegilops tauschii* and *Triticum aestivum*: implications for the origins of hexaploid wheats. *Theor. Appl. Genet.* **112,** 1563–1572 (2006).
47. Gu, Y. Q. *et al.* Types and rates of sequence evolution at the high-molecular-weight glutenin locus in hexaploid wheat and its ancestral genomes. *Genetics* **174,** 1493–1504 (2006).
48. Colledge, S. & Conolly, J. (eds) *The Origins and Spread of Domestic Plants in Southwest Asia and Europe* (Left Coast, 2006).
49. Kilian, B. *et al.* Independent wheat B and G genome origins in outcrossing *Aegilops* progenitor haplotypes. *Mol. Biol. Evol.* **24,** 217–227 (2007).
50. Skoglund, P. Diet, cooking and cosmology: interpreting the evidence of Bronze Age plant macrofossils. *Curr. Swed. Archaeol.* **7,** 149–160 (1999).
51. Anderson, E. A. & Williams, L. Maize and sorghum as a mixed crop in Honduras. *Ann. Mo. Bot. Gard.* **41,** 213–221 (1954).
52. Bradbury, L. M. T., Henry, R. J., Jin, Q., Reinke, R. F. & Waters, D. L. E. A perfect marker for fragrance genotyping in rice. *Mol. Breed.* **16,** 279–283 (2005).
53. Sakamoto, S. in *Redefining Nature: Ecology, Culture and Domestication* (eds Ellen, R. & Fujui, K.) 215–231 (Berg, 1996).
54. Yoshida, S. in *Vegeculture in Eastern Asia and Oceania* (eds Yoshida, S. & Matthews, P.) 31–44 (Japan Center for Area Studies, National Museum of Ethnology, Osaka, 2002).
55. Hirano, H. Y., Eiguchi, M. & Sano, Y. A single base change altered the regulation of the *waxy* gene at the posttranscriptional level during the domestication of rice. *Mol. Biol. Evol.* **15,** 978–987 (1998).
56. Olsen, K. M. & Purugganan, M. D. Molecular evidence on the origin and evolution of glutinous rice. *Genetics* **162,** 941–950 (2002).
57. Olsen, K. M. *et al.* Selection under domestication: evidence for a sweep in the rice *waxy* genomic region. *Genetics* **173,** 975–983 (2006).
This paper reports the characterization of a selective sweep in the *waxy* gene, which underlies the origin and spread of sticky rice.
58. Fukunaga, K., Kawase, M. & Kato, K. Structural variation in the *waxy* gene and differentiation in foxtail millet (*Setaria italica* (L.) P. Beauv.): implications for multiple origins of the waxy phenotype. *Mol. Genet. Genomics* **268,** 214–222 (2002).
59. Fuller, D. Agricultural origins and frontiers in South Asia: a working synthesis. *J. World Prehist.* **20,** 1–86 (2006).
60. Bogaard, A. *Neolithic Farming in Central Europe: An Archaeobotanical Study of Crop Husbandry Practices C5500–2200 BC* (Routledge, 2004).
61. Kreuz, A., Marinova, E., Schafer, E. & Wiethold, J. A comparison of Early Neolithic crop and weed assemblages from the Linearbankeramik and the Bulgarian Neolithic culture: differences and similarities. *Veg. Hist. Archaeobot.* **14,** 237–258 (2005).
62. Colledge, S., Conolly, J. & Shennan, S. The evolution of Neolithic farming from SW Asian origins to NW European limits. *Eur. J. Archaeol.* **8,** 137–156 (2005).
This paper provides a quantitative analysis of crop assemblages and shows the role of selection and drift in the creation of a suitable crop package for new environments.
63. Paterson, A. H. What has QTL mapping taught us about plant domestication? *New Phytol.* **154,** 591–608 (2002).
64. Willis, D. M. & Burke, J. Quantitative trait locus analysis of the early domestication of sunflower. *Genetics* **176,** 2589–2599 (2007).
65. Li, W. & Gill, B. S. Multiple genetic pathways for seed shattering in the grasses. *Funct. Integr. Genomics* **6,** 300–309 (2006).
66. Cai, H. W. & Morishima, H. QTL clusters reflect character associations in wild and cultivated rice. *Theor. Appl. Genet.* **104,** 1217–1228 (2003).
67. Peng, J. *et al.* Domestication quantitative trait loci in *Triticum dicoccoides*, the progenitor of wheat. *Proc. Natl Acad. Sci. USA* **100,** 2489–2494 (2003).
68. Fan. C. *et al. GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theor. Appl. Genet.* **112,** 1164–1171 (2006).
69. Prethepha, P. The fragrance (*fgr*) gene in natural populations of wild rice (*Oryza rufipogon* Griff.). *Genet. Resour. Crop Evol.* **56,** 13–18 (2009).
70. Wang, R. L., Stec, A., Hey, J., Lukens, L. & Doebley, J. The limits of selection during maize domestication. *Nature* **398,** 236–239 (1999).
This paper describes the cloning of the maize *tb1* gene, the first domestication gene cloned in a cereal crop, and discusses a selective sweep at the promoter region.
71. Maynard-Smith, J. & Haigh, J. The hitchhiking effect of a favorable gene. *Genet. Res.* **23,** 23–35 (1974).
72. Clark, R. M., Linton, E., Messing, J. & Doebley, J. F. Pattern of diversity in the genomic region near the maize domestication gene *tb1*. *Proc. Natl Acad. Sci. USA* **101,** 700–707 (2004).
73. Palaisa, K., Morgante, M., Tingey, S. & Rafalski, A. Long-range patterns of diversity and linkage disequilibrium surrounding the maize *Y1* gene are indicative of an asymmetric selective sweep. *Proc. Natl Acad. Sci. USA* **101,** 9885–9890 (2004).
74. Sweeney, M. T. *et al.* Global dissemination of a single mutation conferring white pericarp in rice. *PLoS Genet.* **3,** 1418–1424 (2007).

Acknowledgements We thank K. Olsen and S. Colledge for critical reading of the manuscript. Work in the Purugganan laboratory is funded in part by a grant from the National Science Foundation Plant Genome Research Program.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence should be addressed to M.D.P. (mp132@nyu.edu).

ETS rearrangements and prostate cancer initiation

Arising from: Tomlins *et al*. *Nature* **448,** 595–599 (2007)

The first recurrent translocation event in prostate cancer has been recently described[1]; it results in the translocation of an ETS (E26 transformation specific) transcription factor (*ERG* or *ETV1*) to the *TMPRSS2* promoter region, which contains androgen responsive elements[1]. The *TMPRSS2:ERG* genetic rearrangement has been reported to occur in approximately 40% of primary prostate tumours (*ETV1* genetic rearrangements occur at a much lower frequency), and it results in the aberrant androgen-regulated expression of ERG[1–3]. Tomlins *et al.*[4] concluded that ETS genetic rearrangements are sufficient to initiate prostate neoplasia. However, here we show that ETS genetic rearrangements may in fact represent progression events rather than initiation events in prostate tumorigenesis. To this end, we demonstrate that the prostate-specific overexpression of ERG does not initiate prostate tumorigenesis.

We have found that mice overexpressing ERG (expression confirmed by quantitative PCR with reverse transcription (qRT–PCR), western blotting and immunohistochemistry) under the control of the probasin promoter (ARR2Pb, B6J background strain) do not develop neoplasia, and show only a very subtle phenotype of nuclear atypia (prominent nucleoli) without an increase in cellular proliferation (Fig. 1a). This is similar to what was shown by Tomlins *et al.* in their description of *ETV1* transgenic mice[4]. These subtle nuclear changes without an increase in cellular layers are not sufficient to be classified as prostatic intraepithelial neoplasia (PIN), and are also frequently observed in the wild-type mouse prostate (Fig. 1a). We observe no notable differences in the prostate phenotype across all prostatic lobes (anterior, ventral and dorsal-lateral) between *ERG* transgenic and wild-type littermate mice. The subtle histological changes observed in *ETV1* and *ERG* transgenic mice are markedly different from human high-grade PIN (HGPIN) and from the HGPIN lesions that develop in *Pten* heterozygous mice (Fig. 1b). Furthermore, we did not observe an increase in the proliferative rate in the prostates of mice overexpressing *ERG* compared to wild-type controls. As measured by Ki67 staining, on average 1% of the prostate epithelial cells in both ARR2Pb-*ERG* and wild-type mice were positive (Fig. 1c). With analyses at up to 18 months of age, no ARR2Pb-*ERG* mice have shown any change in phenotype. Whereas Tomlins *et al.* concluded that ETS genetic rearrangements are sufficient to initiate prostate neoplasia, we present data to suggest that ETS genetic rearrangements may in fact represent progression events rather than initiation events in prostate tumorigenesis, as there are no proliferative and pathological changes consistent with HGPIN found in either *ERG* or *ETV1* mice.

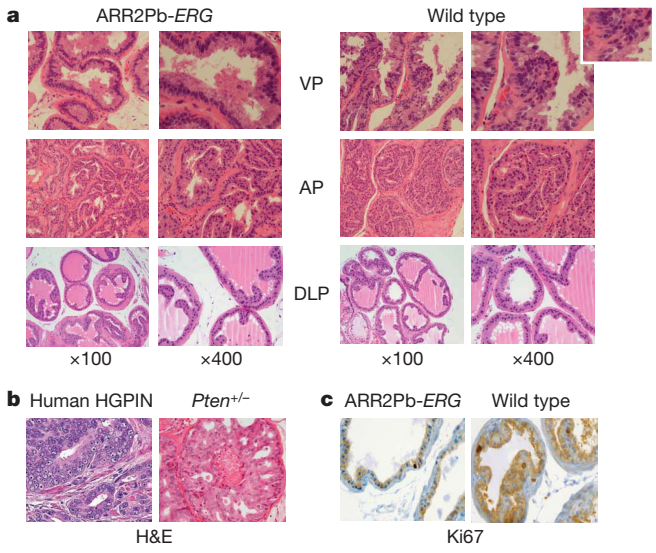


Figure 1 | Prostate specific overexpression of ERG does not induce high grade prostatic intra-epithelial neoplasia. a, A total of 24 wild-type and 24 *ERG* transgenic mice were phenotypically characterized from one founding line after the establishment that three independent founding lines produced a similar phenotype. Low-power (×100) and high-power (×400) representative sections are shown for mice 6 months of age, demonstrating prominent nucleoli in wild-type mouse prostate glands. AP, anterior prostate lobes; DLP, dorsal-lateral prostate lobes; VP, ventral prostate lobes. **b**, Representative histology of human HGPIN (left) and HGPIN in 12-month-old *Pten* heterozygous mice (right). H&E, haematoxylin and eosin. **c**, Immunohistochemistry demonstrated no difference in Ki67 staining between wild-type and *ERG* transgenic mice. Original magnification in **b** and **c**, ×400.

Furthermore, the *ERG* translocation is infrequently found in human HGPIN and only in a minority (approximately 10–20%) of patients who also have the translocation present in associated adenocarcinoma of the prostate. Most prostate cancer specimens with *ERG* genetic rearrangements do not show this rearrangement in the associated HGPIN. Therefore, the *TMPRSS2:ERG* translocation seems to be an early event in human prostate tumorigenesis, but one associated with progression from HGPIN to cancer.

Using mouse modelling, we have demonstrated that the aberrant expression of ERG is not sufficient to initiate neoplastic transformation but instead may cooperate with other genetic events to promote prostate cancer progression. We propose a working model whereby genetic initiating events conferring a proliferative advantage select for cooperating ETS genetic rearrangements that promote an invasive phenotype.

METHODS

Mice (B6J background strain) expressing *ERG* under the control of the probasin promoter (ARR2Pb) were generated, genotyped and examined for transgene expression by qRT–PCR, western blotting and immunohistochemistry. We generated and analysed three independent *ERG* lines. Founders were subsequently bred and four mice of each genotype were euthanized at 2, 4, 6, 8, 12 and 18 months of age. Prostate tissues were procured for formalin fixation, paraffin embedding and frozen storage for future molecular analyses.

Brett S. Carver[1,3], Jennifer Tran[1], Zhenbang Chen[1,5], Arkaitz Carracedo-Perez[1], Andrea Alimonti[1,5], Caterina Nardella[1,5], Anuradha Gopalan[2], Peter T. Scardino[3], Carlos Cordon-Cardo[4], William Gerald[2] & Pier Paolo Pandolfi[1,2,5]

[1]Cancer Biology and Genetics Program, Sloan-Kettering Institute, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA.
e-mail: ppandolf@bidmc.harvard.edu
[2]Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA.
[3]Department of Surgery, Division of Urology, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA.
[4]Department of Pathology, Columbia University, New York, New York 10032, USA.
[5]Cancer Genetics Program, Beth Israel Deaconess Cancer Center and Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02215, USA.

Received 19 March; accepted 20 November 2008.

1. Tomlins, S. A. *et al.* Recurrent fusion of *TMPRSS2* and ETS transcription factor genes in prostate cancer. *Science* **310,** 644–648 (2005).
2. Tu J. J., Rohan S., Kao J., Kitabayashi N., Mathew S. & Chen Y.T.. Gene fusions between *TMPRSS2* and *ETS* family genes in prostate cancer: frequency and transcript variant analysis by RT–PCR and FISH on paraffin-embedded tissues. *Mod. Pathol.* **20,** 921–928 (2007).
3. Perner, S. *et al.* TMPRSS2-ERG fusion prostate cancer: an early molecular event associated with invasion. *Am. J. Surg. Pathol.* **31,** 882–888 (2007).
4. Tomlins, S. A. *et al.* Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. *Nature* **448,** 595–599 (2007).

doi:10.1038/nature07738

Tomlins et al. *reply*

Replying to: Carver, B. S. *et al. Nature* **457**, doi:10.1038/nature07738 (2009)

Carver *et al.*[1] question our recent report that mice expressing *ETV1* under the control of the probasin promoter (ARR2Pb) develop mouse prostatic intraepithelial neoplasia (mPIN)[2]. They report the generation of transgenic ARR2Pb-*ERG* mice with no phenotypic differences from control mice. They propose that this demonstrates that ETS genetic rearrangements do not initiate prostate tumorigenesis and use data from human prostate cancer studies to propose that ETS rearrangements are associated with progression from PIN to prostate cancer. Although we and others have shown that ARR2Pb-*ETV1* and ARR2Pb-*ERG* mice develop mPIN, we have consistently proposed that in human prostate cancer development, ETS rearrangements mediate the transition from PIN to cancer.

Our blinded histopathological evaluation of the ARR2Pb-*ETV1* mice included two genitourinary pathologists (R.B.S. and M.A.R.), one of whom is a co-author on 'The consensus report from the Bar Harbor Meeting of the Mouse Models of Human Cancer Consortium Prostate Pathology Committee' (M.A.R.), and the diagnosis of mPIN without progression to carcinoma was made according to those criteria[3]. Notably, members of the Bar Harbour Committee observed a wide spectrum of morphological alterations that were all considered mPIN, and the exact replication of human high-grade PIN is not required to define mPIN. In fact, the group of human and animal pathologists adopted this view to reflect the wide spectrum of variations observed between different mouse models of prostate cancer.

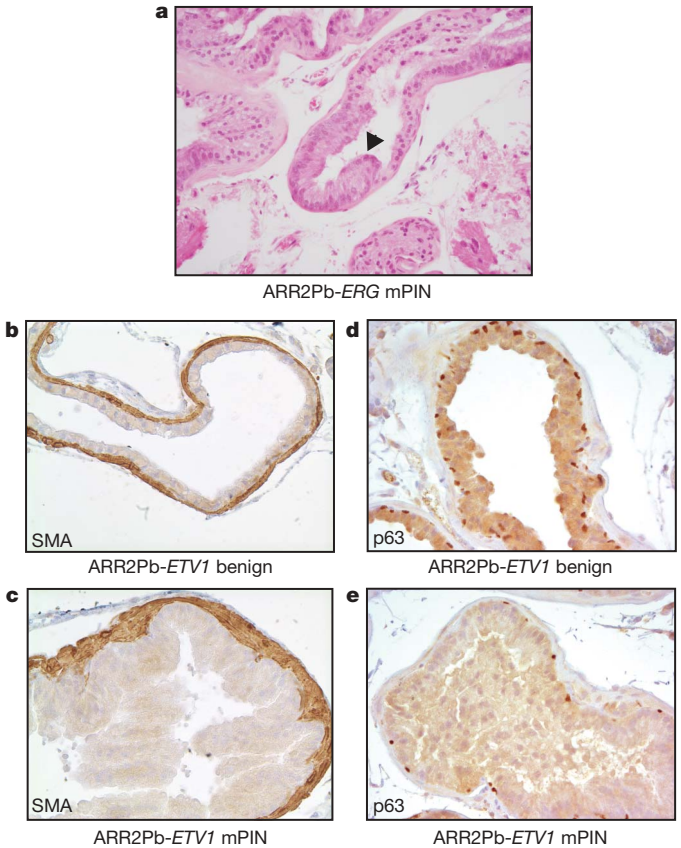


Figure 1 | mPIN in ARR2Pb-*ERG* and ARR2Pb-*ETV1* mice. a, In ARR2Pb-*ERG* mice, distinct areas of proliferation in sporadic glands, consistent with the definition of mPIN (black arrowhead), were observed adjacent to normal prostatic epithelium. **b**–**e**, Loss of the circumferential basal layer in mPIN lesions in ARR2Pb-*ETV1* mice. Consistent with the focal nature of mPIN, normal areas and mPIN were observed in the prostate of ARR2Pb-*ETV1* mice. Immunohistochemistry with smooth muscle actin (SMA) demonstrates a continuous fibromuscular layer around benign glands (**b**) and all mPIN lesions (**c**), whereas the basal cell marker p63 demonstrates the loss of circumferential basal cells in mPIN foci (**e**) compared to normal glands (**d**) in the dorsolateral prostate of a ARR2Pb-*ETV1* mouse. Original magnification for all images is ×400.

We also generated ARR2Pb-*ERG* mice, which again by the Bar Harbour Committee classification develop mPIN without progression to carcinoma (Fig. 1a). Similar to ARR2Pb-*ETV1* mice, ARR2Pb-*ERG* mice have focal lesions showing nuclear atypia, including stratification, hyperchromasia and macronucleoli[4]. The development of mPIN without progression to carcinoma in ARR2Pb-*ERG* transgenic mice has also been previously described[5]. In this model, a decrease in basal epithelial cells was shown in mPIN lesions, and luminal epithelial cells directly contacted the stromal cell compartment[5]. We found a similar loss of the circumferential basal epithelial layer in mPIN lesions from our ARR2Pb-*ETV1* (Fig. 1b–e) and ARR2Pb-*ERG* mice[4], which is a hallmark of prostate carcinoma development in both mice and humans[6]. As our ARR2Pb-*ERG* mice and those used by Klezovitch *et al.*[5] were generated on different backgrounds, this strongly supports a phenotypic effect in ARR2Pb-*ERG* mice. It is unclear whether Carver *et al.* looked for changes in the relationship between the basal epithelial layer and the stromal compartment in their model.

Although we feel that prostate specific expression of *ERG* or *ETV1* induces PIN in mice, we have never claimed that this (or any other evidence) supports ETS rearrangements initiating human prostate cancer tumorigenesis through the development of PIN. Instead, through our results including *ERG* expression in human PIN and prostate cancer by DNA microarray analysis[7], fluorescence *in situ* hybridization (FISH) data showing the frequency of ETS rearrangements in PIN and prostate cancer[8,9], *ERG* knockdown in VCaP (a *TMPRSS2:ERG* positive prostate cancer cell line) deregulating the transcriptional program differentiating PIN and prostate cancer[4], *in vitro* data demonstrating a role for *ETV1* and *ERG* in invasion[2,4], and ARR2Pb-*ERG* and ARR2Pb-*ETV1* mice not developing frank carcinoma[2,4], we have consistently proposed that ETS gene fusions in humans mediate the transition from prostate cells with pre-existing lesions (such as cells in PIN foci) to carcinoma[2,4,7–10]. We feel that models combining ETS rearrangements and other early lesions in prostate cancer development have the potential to transform *in vitro* and *in vivo* prostate cancer research.

Scott A. Tomlins[1], Bharathi Laxman[1], Saravana M. Dhanasekaran[1], Beth E. Helgeson[1], Xuhong Cao[1], David S. Morris[2], Anjana Menon[1], Xiaojun Jing[1], Qi Cao[1], Bo Han[1], Jindan Yu[1], Lei Wang[1], James E. Montie[2,4], Mark A. Rubin[5,6]†, Kenneth J. Pienta[2,3,4], Diane Roulston[1], Rajal B. Shah[1,2,4], Sooryanarayana Varambally[1,4], Rohit Mehra[1,4] & Arul M. Chinnaiyan[1,2,4]

[1]Michigan Center for Translational Pathology, Department of Pathology, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA.
e-mail: arul@umich.edu
[2]Department of Urology, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA.
[3]Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA.
[4]Comprehensive Cancer Center, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA.
[5]Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, 02115, USA.
[6]Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts, 02115, USA.
†Present address: Department of Pathology and Laboratory Medicine, Cornell University, New York, New York 10021, USA.

1. Carver, B. S. *et al.* ETS rearrangements and prostate cancer initiation. *Nature* **457,** doi:10.1038/nature07738 (2009).
2. Tomlins, S. A. *et al.* Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. *Nature* **448,** 595–599 (2007).
3. Shappell, S. B. *et al.* Prostate pathology of genetically engineered mice: definitions and classification. The consensus report from the Bar Harbor meeting of the Mouse Models of Human Cancer Consortium Prostate Pathology Committee. *Cancer Res.* **64,** 2270–2305 (2004).
4. Tomlins, S. A. *et al.* Role of the *TMPRSS2-ERG* gene fusion in prostate cancer. *Neoplasia* **10,** 177–188 (2008).
5. Klezovitch, O. *et al.* A causal role for ERG in neoplastic transformation of prostate epithelium. *Proc. Natl Acad. Sci. USA* **105,** 2105–2110 (2008).
6. DeMarzo, A. M., Nelson, W. G., Isaacs, W. B. & Epstein, J. I. Pathological and molecular aspects of prostate cancer. *Lancet* **361,** 955–964 (2003).
7. Tomlins, S. A. *et al.* Integrative molecular concept modeling of prostate cancer progression. *Nature Genet.* **39,** 41–51 (2007).
8. Mosquera, J. M. *et al.* Characterization of *TMPRSS2-ERG* fusion high-grade prostatic intraepithelial neoplasia and potential clinical implications. *Clin. Cancer Res.* **14,** 3380–3385 (2008).
9. Perner, S. *et al.* TMPRSS2-ERG fusion prostate cancer: an early molecular event associated with invasion. *Am. J. Surg. Pathol.* **31,** 882–888 (2007).
10. Kumar-Sinha, C., Tomlins, S. A. & Chinnaiyan, A. M. Recurrent gene fusions in prostate cancer. *Nature Rev. Cancer* **8,** 497–511 (2008).

doi:10.1038/nature07739

ARTICLES

Life without a wall or division machine in *Bacillus subtilis*

M. Leaver[1], P. Domínguez-Cuevas[1], J. M. Coxhead[2], R. A. Daniel[1] & J. Errington[1]

The cell wall is an essential structure for virtually all bacteria, forming a tough outer shell that protects the cell from damage and osmotic lysis. It is the target of our best antibiotics. L-form strains are wall-deficient derivatives of common bacteria that have been studied for decades. However, they are difficult to generate and typically require growth for many generations on osmotically protective media with antibiotics or enzymes that kill walled forms. Despite their potential importance for understanding antibiotic resistance and pathogenesis, little is known about their basic cell biology or their means of propagation. We have developed a controllable system for generating L-forms in the highly tractable model bacterium *Bacillus subtilis*. Here, using genome sequencing, we identify a single point mutation that predisposes cells to grow without a wall. We show that propagation of L-forms does not require the normal FtsZ-dependent division machine but occurs by a remarkable extrusion-resolution mechanism. This novel form of propagation provides insights into how early forms of cellular life may have proliferated.

Binary fission is a widely conserved mechanism required for the proliferation of almost all cells. In bacteria, the central protein in cytokinesis is the bacterial tubulin homologue, FtsZ (filamentous temperature sensitive protein Z), which forms a ring (the Z-ring) to which other proteins are recruited[1]. The division machine then constricts in parallel with synthesis of a cross wall that matures into the new polar caps of the daughter cells[2] (Supplementary Fig. 1a). The machine includes enzymes called penicillin-binding proteins, which synthesize peptidoglycan, the major component of the wall[3] (Supplementary Fig. 1b). L-form bacteria[4,5] can proliferate but are thought not to have any wall; they offer an interesting model for investigating the details of Z-ring function[6], particularly whether constriction of the Z-ring requires wall synthesis.

A rapid method for generating L-forms

Existing L-form bacteria are difficult to manipulate[7], and modern genetic tools such as inducible promoters and green fluorescent protein (GFP) fusions have not been used in them. We therefore developed an approach to reproducibly generate stable *B. subtilis* L-forms.

We used strain M96 in which expression of the *murE* operon, which encodes several enzymes essential for the synthesis of the precursor of peptidoglycan[8], is controlled by the xylose inducible promoter P_{xyl}. In the presence of inducer, both the growth rate and the shape of this strain were normal (Fig. 1a–c). In the absence of inducer, cell growth was arrested and cells bulged and lysed (Fig. 1a, d). As previous work has shown that many shape mutants can be rescued by high concentrations of Mg^{2+} or osmoprotectants[9–12], we tested whether these conditions could restore growth to the MurE depletion strain. In the presence of high concentrations of Mg^{2+}, the growth of this strain was restored to a limited extent (Fig. 1a). Phase contrast microscopy of cells incubated in medium containing Mg^{2+} and sucrose revealed that lysis was reduced, and large, amorphous cells accumulated (Fig. 1e) that were similar in appearance to classically derived[7,13,14] and antibiotic induced L-forms[6]. These cells appeared to have little or no wall: when squashed between a microscope slide and a cover slip they packed together in a tessellated pattern, showing not only that they had lost their shape but also the rigidity of their envelope (Fig. 1f).

Unfortunately, these cells did not undergo sustained proliferation and the cultures were rapidly overrun with rod-shaped cells (Fig. 1g, arrowhead) presumably derived by spontaneous mutations in *xylR* or P_{xyl} that resulted in constitutive expression of *murE* in the absence of inducer. By introducing a second copy of the *xylR* repressor gene and selection with high concentrations of penicillin G (PenG, an antibiotic that inhibits cell wall synthesis by inactivating penicillin-binding proteins), occasional colonies of pure L-form-like cells were obtained, which could be propagated indefinitely (Methods). The frequency at which L-form-like colonies were generated was lower than the rate of formation of xylose resistant rod-shaped mutants, which presumably arose by mutation of the P_{xyl} promoter or double *xylR* mutants. It therefore seems likely that one or more secondary mutations are needed to produce cells that can proliferate in this L-form-like state.

A strain generated by this method (Bs115 sup21) had many of the properties of L-forms made by classical methods (Fig. 1h, Supplementary Results and Discussion)[7,15]. Importantly, strain Bs115 sup21 was about 1,000 times more resistant to PenG than the wild-type strain (Supplementary Fig. 2), consistent with the idea that L-forms do not make peptidoglycan (PenG probably has non-specific toxic effects at very high concentrations). Strain Bs115 sup21 did not require PenG to grow in the L-form state and, in the absence of PenG and presence of *murE* inducer (xylose), only reverted to rod shape at low frequency. This suggests that at least one of the putative secondary mutations that it had acquired prevents it from making cell wall.

A mutation that predisposes for L-form proliferation

The frequency with which L-forms arose suggested that one or two mutations were needed to generate L-forms. To test this, we performed shotgun sequencing[16] of the complete genome of a freshly derived L-form strain (Bs115 sup23). Approximately 700 polymorphisms were identified by comparison of the sequence to that published in the SubtiList database[17]. Only four of these were also found to be different

[1]Institute for Cell and Molecular Biosciences, Newcastle University, Framlington Place, Newcastle Upon Tyne NE2 4HH, UK. [2]Institute for Human Genetics, Newcastle University, International Centre for Life, Central Parkway, Newcastle Upon Tyne NE1 3BZ, UK.

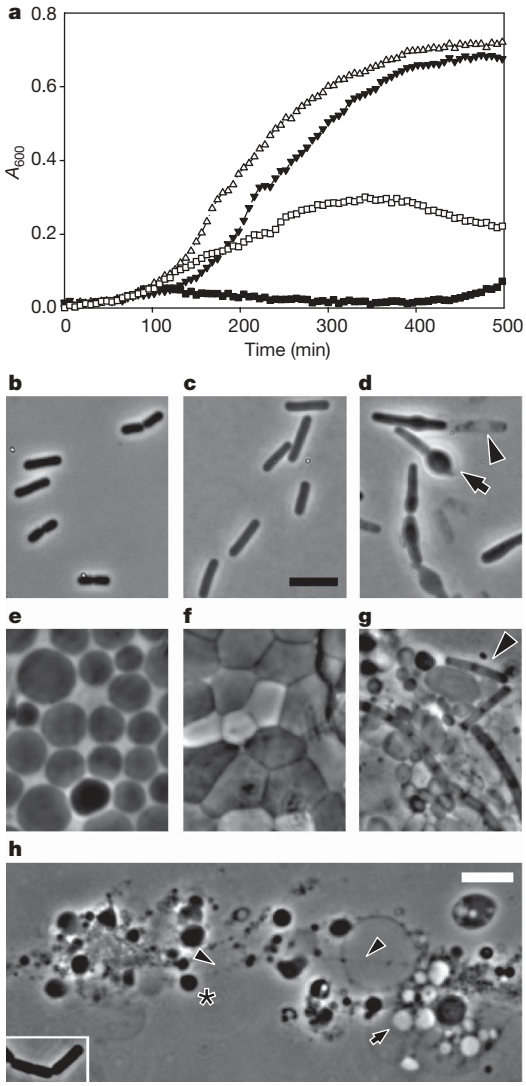


Figure 1 | Controlled generation of *B. subtilis* L-forms in cells with a repressible *murE* operon. **a**, The growth of the MurE depletion strain (M96) grown with *murE* expression (upright open triangle) and without *murE* expression either with (open square) or without Mg^{2+} (filled square) compared to the wild-type strain (inverted filled triangle). **b**–**h**, Phase contrast images of wild-type strain (**b**) and strain M96 expressing *murE* (**c**). Cells of strain M96 not expressing *murE* are bulged (**d**, arrow) and lytic (**d**, arrowhead). Cells of strain M96 not expressing *murE* form large spheroids in the presence of Mg^{2+} and sucrose (**e**) and tessellate when forced together (**f**). Cells of strain M96 not expressing *murE* grown in the presence of Mg^{2+}, sucrose and sloppy agar (**g**) have a similar appearance to L-form cells but can revert to rod-shaped cells (arrowhead). L-form strain Bs115 sup21 grown in liquid medium supplemented with $MgCl_2$, sucrose and penicillin G (**h**) are spherical and often connected by thin strands (arrow heads). Cells appeared either phase bright (arrow) or phase dark (asterisk). Inset, wild-type rods for comparison. Scale bar, 5 μm.

from a recently re-sequenced *B. subtilis* strain (A. Danchin, personal communication). Three of these four mutations turned out to be identical to substitutions in the immediate parent strain, Bs115. The one remaining polymorphism would generate a single amino acid substitution, D92E (GAT codon to GAA), in the *yqiD* gene. This gene encodes a protein of 272 amino acids, which is 45% identical to IspA of *Escherichia coli* (Supplementary Fig. 3). IspA catalyses the formation of farnesyl pyrophosphate in the isoprenoid synthetic pathway[18]. This pathway leads to the formation of several essential lipids, including bactoprenol, involved in peptidoglycan synthesis (and teichoic acids in *B. subtilis*). The D92E substitution (corresponding to D90 in the *E. coli* homologue, Supplementary Fig. 3) affects a conserved residue in the substrate-binding cavity of the enzyme[19] (Supplementary Fig. 4a, b) which is required for enzyme activity[20].

To test whether the mutation was related to the L-form phenotype, we constructed a strain with an extra copy of the mutant form of *ispA* driven by the IPTG (isopropyl-β-D-thiogalactoside)-dependent P_{spac} promoter. In rod-shaped cells, the integration did not significantly affect either growth rate or cell morphology, either in the presence or absence of IPTG (data not shown). We used the selective regime described above to test for L-form conversion in the presence of IPTG to express the mutant copy of *ispA* along with the wild-type copy. The results revealed about a 1,000-fold increase in the frequency of L-form-colony formation (average over three experiments, Supplementary Fig. 4c). In the typical experiment shown, the parent strain gave only three or four colonies when about a million cells were plated, whereas the mutant strain gave about 50 colonies from a 100-fold more dilute sample.

These results showed that a single point mutation in the *ispA* gene is sufficient to enable cells to grow without a wall. Although we do not yet understand the mechanism, it can be used to generate L-form strains at will, opening up a new era of research on these interesting cells and perhaps helping understanding of their role in pathogenesis and antibiotic resistance[21].

FtsZ-independent division

Although early biochemical experiments indicated that stable L-forms of Gram-positive bacteria do not synthesize detectable amounts of peptidoglycan[22–24], D'Ari and co-workers have recently reported that *E. coli* L-forms require residual peptidoglycan for division[6]. To test if this was also the case in our L-forms, we constructed a strain (Bs126) that was unable to synthesize D-alanine, which is an essential component of peptidoglycan. Rod-shaped bacteria required an exogenous source of D-alanine for growth, but an L-form strain derived from Bs126 grew without D-alanine (Supplementary Results and Discussion), suggesting that L-forms do not require peptidoglycan for division.

To begin investigating how L-forms divide without a cell wall, we constructed a strain that expressed an FtsZ–GFP fusion (strain Bs119 sup4). Fluorescence microscopy revealed spots and arcs of FtsZ–GFP, as well as complex networks of polymers (Fig. 2a). Unexpectedly, rings of FtsZ that might be supporting division were not evident. In the light of this observation, we tested whether FtsZ was actually required for L-form proliferation. To do so, we first used a strain in which the only copy of *ftsZ* was controlled by P_{spac}. As previously reported[25], when FtsZ was depleted in wild-type cells (strain BB11), division was completely inhibited and cells elongated for a few generations before lysing. Remarkably, when FtsZ was depleted in L-form strain Bs120 sup20, the growth rate was unaffected (Fig. 2b) and the strain could be indefinitely propagated without FtsZ expression. Western blot analysis confirmed that, when grown in the absence of inducer, FtsZ levels were barely detectable (Fig. 2b inset, lane 3).

To test if this residual amount of FtsZ was required for growth of the L-form, we attempted to make a strain with a complete deletion of *ftsZ* and *ftsA*, which form a two gene operon. The chromosomal copy of *ftsAZ* was deleted in the presence of an ectopic copy of the operon on the unstable plasmid, pLOSS[26] (Fig. 2c, lanes 5–7, Supplementary Results and Discussion). As FtsZ is essential in wild-type cells, only cells that keep the plasmid should survive. The presence or absence of pLOSS was assayed by PCR analysis and blue/white screening: colonies of plasmid containing cells appear blue on plates containing the β-galactosidase substrate X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside), because pLOSS expresses *lacZ* (encoding β-galactosidase; Supplementary Fig. 6a). As expected, in rod-shaped bacteria with *ftsAZ* deleted, all colonies were blue, demonstrating the requirement for pLOSS::*ftsAZ* (Supplementary Fig. 6c). However, an L-form strain with the same genotype gave only white colonies, suggesting that the strain had lost the plasmid and therefore that L-forms do not require FtsZ. Conventional PCR and quantitative

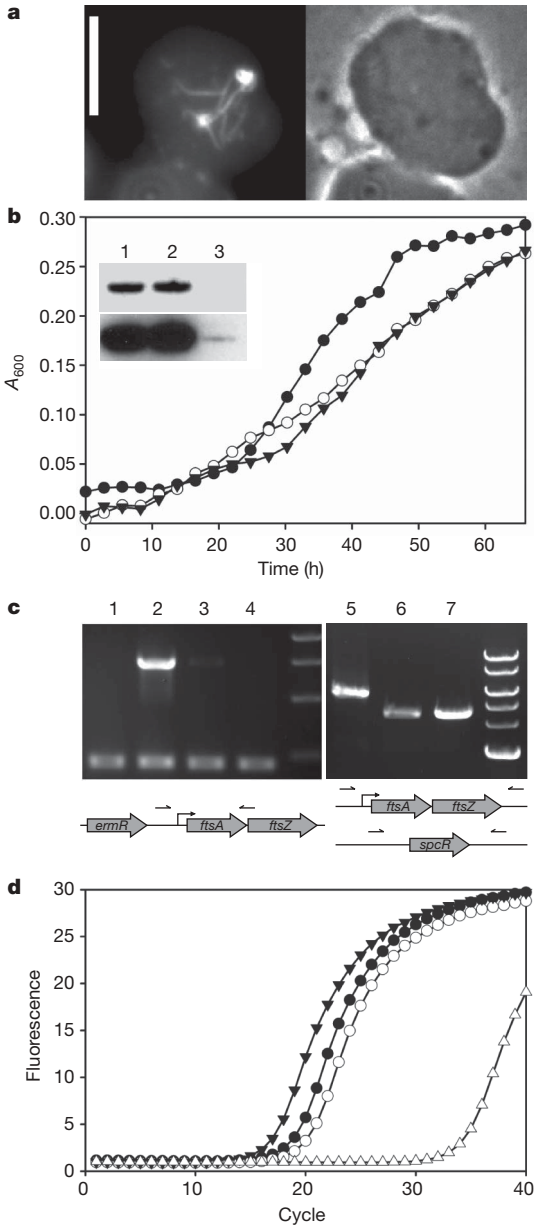


Figure 2 | L-forms do not require FtsZ for division. **a**, Localization of GFP–FtsZ in an L-form (strain Bs119 sup4). Left panel, GFP fluorescence; right panel, phase contrast image. Scale bar, 5 μm. **b**, Growth of the *ftsZ*$^+$ L-form (Bs115 sup21, filled circles) and the FtsZ depletion strain (Bs120 sup20) with (open circles) and without FtsZ expression (filled inverted triangles). Inset, western blot showing levels of FtsZ in *ftsZ*$^+$ L-forms (lane 1), from the FtsZ depletion strain with and without expression of FtsZ (lanes 2 and 3). Top panel, 15 s exposure; bottom panel, 5 min exposure. **c**, **d**, Control experiments showing that plasmid pLOSS::*ftsAZ* is lost in an L-form strain with a deletion of *ftsAZ* but not in a similar rod-shaped strain. **c**, Lane 1, PCR with oligonucleotides that bind on the plasmid backbone and in *ftsA* (diagram) amplified a 1.5 kb product from strain Bs161 (rod-form *ftsAZ*::*spc* pLOSS::*ftsAZ*, lane 2) but not from the wild-type strain (168, lane 1). A weak PCR product was amplified from chromosomal DNA of an L-form derived from Bs161 (lane 3) but no product was detected after this strain was plated for single colonies (lane 4). A control PCR reaction gave a 500 bp fragment from all chromosomal DNAs. PCR with oligonucleotides that bind either side of the *ftsAZ* operon (diagram) gave a 6 kb product with DNA from wild-type *B. subtilis* (lane 5) but only a shorter product with DNA from strain Bs161 or the L-form derived from Bs161 (lanes 6 and 7). **d**, Quantitative real-time PCR analysis of wild-type *B. subtilis* (circles) and the L-form strain derived from strain Bs161 (triangles). Panel shows the accumulation of PCR product (SYBR green fluorescence) during a 40 cycle PCR amplifying markers at the origin (filled symbols) and *ftsZ* loci (open symbols).

real time PCR confirmed that most rod-shaped cells kept pLOSS::*ftsAZ* but that in L-form cells the plasmid was either completely lost or present only in a very small sub-population of cells (about 1 plasmid per 250,000 chromosome origins; Fig. 2c, d; Supplementary Results and Discussion). Thus, the 'essential' *ftsZ* gene becomes non-essential in L-forms.

Proliferation by extrusion and resolution

To determine how L-forms proliferate without FtsZ, we developed conditions for long-term time-lapse imaging. L-form strains were grown in a glass-bottom microwell dish on the stage of the microscope at 30 °C (Methods). Observations of numerous fields over periods as long as 12 h (Supplementary Movie 1) showed that, although not all L-forms undergo growth, many cells, across the entire size range, showed a significant increase in size, indicating that they are metabolizing and synthesizing at least some cellular materials.

We observed several events that appeared to represent the normal mode of proliferation of L-form cells (Fig. 3). Initially, the more-or-less round L-form in Fig. 3a (Supplementary Movie 2) increased in size (arrowed, 10–320 min). It then went through a period when a series of shape perturbations occurred, characterized by the formation of transient blunt protrusions (370–410 min). These appeared and then retracted, and re-emerged from different points on the cell surface. Finally, a pseudopodium-like protrusion emerged and elongated over a period of about 70 min (Fig. 3a, arrowhead). This protrusion then resolved into about five more-or-less round bodies (540–560 min) that we believe represent progeny (see below). Although extrusion was slow, resolution can be rapid: in a separate experiment, a cell that had formed an extrusion was observed to resolve into three roughly spherical cells in less than one minute (Fig. 3d).

A similar extrusion-resolution process was also evident in smaller L-forms (<2 μm diameter); this sometimes gave rise to progeny of roughly equal size (Fig. 3b, Supplementary Movie 3), similarly to a previous report on division in *B. subtilis* L-forms[15]. This time-lapse imaging also demonstrated that at least some progeny are viable, because the two initial progeny both went on to undertake further division events. Note that one of these subsequent events gave rise to two roughly equal progeny, whereas the other was asymmetrical (Fig. 3b, 430 min, arrows and arrowheads, respectively).

The third kind of event we frequently observed generally involved large L-forms (3.5–6 μm diameter). In the example shown in Fig. 3c (Supplementary Movie 4), an L-form underwent a period of growth as a uniform sphere (10–270 min). This was followed by a period in which multiple, transient bulges occurred on the cell surface (Fig. 3c, 430–590 min, arrowheads). Finally, many small bodies erupted from a few sites on the cell surface and resolved into apparent progeny (Fig. 3c, 710–720 arrows).

Transmission electron micrographs (TEMs) of thin sections of L-forms revealed a heterogeneous population of cells, similar to those of stable L-forms reported previously[13,22]. Importantly, we often observed examples of cells apparently undergoing the proliferation events described above (Fig. 3e–g).

A strain that expressed a GFP-fusion to a DNA binding protein was used as a marker for the nucleoid. Time-lapse microscopy of this strain showed that cell proliferation was accompanied by many extrusion-resolution events (Supplementary Fig. 7). Multiple rounds of division resulted in the formation of a micro-colony, suggesting that progeny were viable. Every large cell-like object in the colony had a GFP signal, showing that the proliferative events are accompanied by segregation of DNA (Supplementary Results and Discussion).

Conclusions and perspective

These experiments were originally prompted by the question of whether Z-ring constriction requires concomitant cell wall synthesis. We noticed that although FtsZ forms polymers in L-forms, it does not form significant numbers of rings, nor is there any sign of FtsZ-associated constriction. The results therefore support a model in which the constriction of the Z-rings is dependent on wall synthesis.

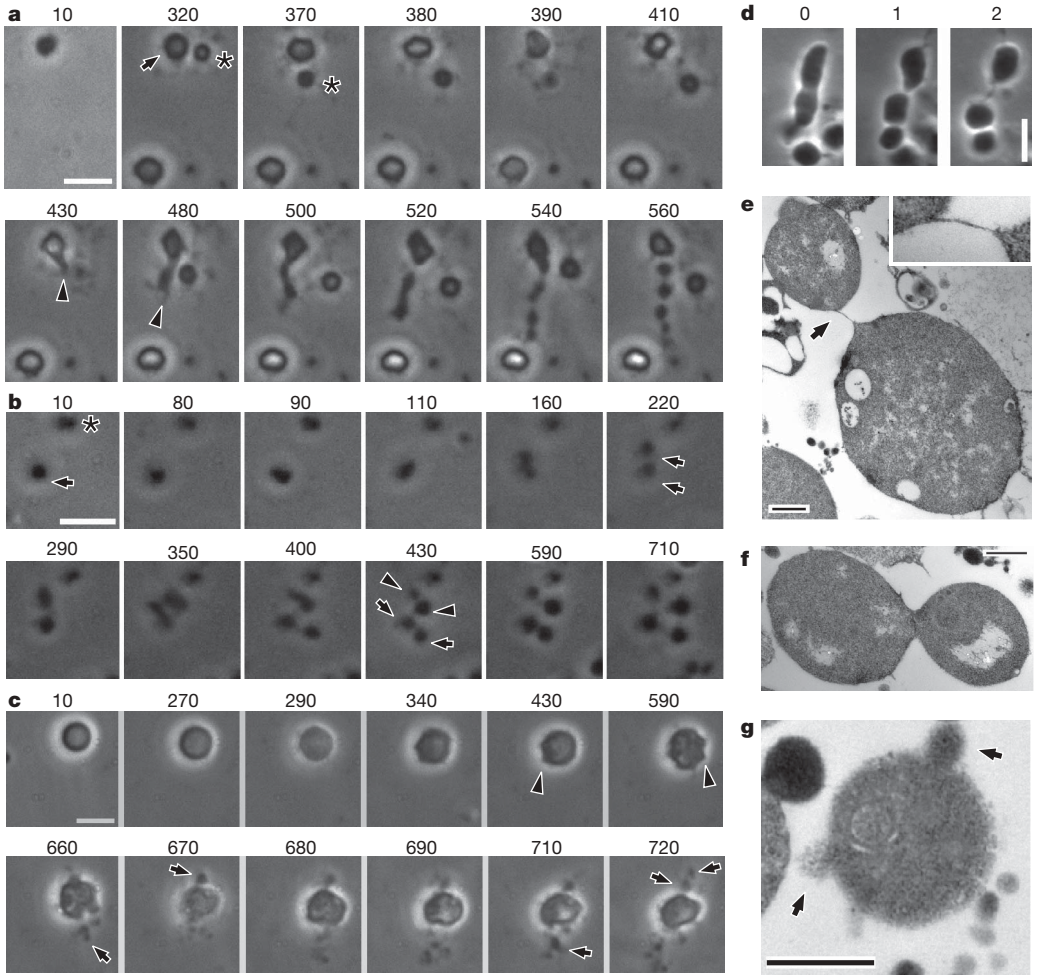


Figure 3 | Novel mechanisms of proliferation revealed by time-lapse imaging of L-forms. **a**–**c**, Still images from Supplementary Movies 2–4, with times indicated in min. Scale bars, 5 μm. **a**, The cell marked with an arrow (at 320 min) begins to form a protrusion at 430 min (arrow head) which resolves into a string of six cells visible at 560 min. Note that another cell (marked with an asterisk at 320 min) drifts into the field of view and was not the result of a proliferative event during the movie. **b**, The cell marked with an arrow (at 10 min) divides after 160 min to produce two cells (220 min, arrows), which divide again after 350 min to give symmetrical progeny (430 min, arrows) and asymmetrical progeny (430 min, arrowheads). Another cell (marked with an asterisk at 10 min) does not divide during the time course. **c**, Surface blebs emerge from the cell (430–590 min, arrowheads), and later, multiple small progeny are extruded (arrows). **d**, An elongated cell at 0 min resolves into three roughly spherical cells after 1 min. **e**–**g**, Transmission electron micrographs of dividing L-forms. **e**, Two cells connected by a thin strand (arrow, and inset at 3× magnification). **f**, A cell apparently at a late stage of a resolution event. **g**, A cell with surface blebs and possible progeny (arrows). Scale bars, 500 nm.

However, it remains possible that the transition to the L-form state may involve other changes that are incompatible with Z-ring constriction. Nevertheless, *ftsZ* could be deleted in L-forms, showing that although it is normally essential in many bacteria, it becomes non-essential in cells growing without a wall.

Prokaryotes that do not have peptidoglycan in their envelopes include members of the class Mollicutes (including the major genera *Mycoplasma*, *Ureaplasma*, *Acholeplasma*, *Spiroplasma* and *Phytoplasma*[27]), the order *Planctomycetales*[28] and the archaea[29]. Of these, some also lack FtsZ, including the *Phytoplasma*, and the archaeal phylum crenarchaeota[30]. Little is known of the replication of these organisms, but the unusual mode of proliferation we have uncovered for *B. subtilis* L-forms may be of relevance to them.

In principle, it is difficult to envisage how extrusion-resolution could occur in the absence of an active force generation system within the cells[31]. Therefore, it seems likely that some kind of cytoskeletal system is involved. FtsZ appears not to be required, so the most likely candidates would be one or more of the MreB (Actin) homologues, of which *B. subtilis* has three: MreB[32], Mbl[33] and MreBH[34]. Alternatively, extrusion could be driven by an active chromosome segregation system. Extrusion of a nucleoid followed by collapse and resealing of the membrane could explain the resolution process we observed. The molecular details underlying this process, reminiscent of pseudopod production in higher cells, will be an important area for future work.

Finally, we speculate that the novel mechanism of proliferation we have discovered has implications for the evolution of early forms of life. Genes for the synthesis of peptidoglycan are found throughout all major branches of the bacterial kingdom[35], suggesting that the last common ancestor already had a wall. However, this cell is likely to have been derived from an earlier wall-less progenitor. We propose that an extrusion-resolution mode of proliferation may have been used in these progenitor cells. Later, when the wall evolved, elements of this mechanism were used to drive segregation along the long axis of the cylindrical tube formed by the rigid wall. The system may have been retained for occasional use when the wall was damaged, for example, during antibiotic 'warfare' between competing microorganisms. The Mollicutes may represent specialized cells that have readopted this mode of replication as part of their adaptation to a life style intimately associated with modern higher eukaryotes.

METHODS SUMMARY

Reproducible method for generation of L-form bacteria. Strain Bs115, or a derivative of that strain, was grown to late exponential phase in rich medium supplemented with 20 mM $MgCl_2$, 0.5 M sucrose and buffered with 20 mM

maleic acid (this supplement is referred to as MSM) and 0.5% w/v xylose. Cells were washed once in MSM in H_2O. 100 µl of cells was added to 5 ml of nutrient broth (NB, Oxoid) with MSM and molten 1% w/v type VII low gelling temperature agarose ('soft agar', Sigma) that had been cooled to 37 °C, then spread on top of a nutrient agar (NA, Oxoid) plate supplemented with MSM, but without xylose. Plates were inverted and incubated at 30 °C overnight in an airtight container. After 12–14 h incubation, 1 cm^2 blocks of sloppy agarose that did not contain colonies of rod-shape cells were spread on fresh NA plates supplemented with MSM and 200 µg ml^{-1} PenG (Sigma). Atypical colonies containing L-form-like cells formed after 3–4 days.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 22 May; accepted 22 December 2008.

1. Rothfield, L., Taghbalout, A. & Shih, Y. L. Spatial control of bacterial division-site placement. *Nature Rev. Microbiol.* **3,** 959–968 (2005).
2. Weiss, D. S. Bacterial cell division and the septal ring. *Mol. Microbiol.* **54,** 588–597 (2004).
3. Höltje, J. V. Growth of the stress-bearing and shape-maintaining murein sacculus of *Escherichia coli. Microbiol. Mol. Biol. Rev.* **62,** 181–203 (1998).
4. Klieneberger, E. The natural occurrence of pleuropneumonia-like organisms in apparent symbiosis with *Streptobacillus moniliformis* and other bacteria. *J. Pathol. Bacteriol.* **40,** 93–105 (1935).
5. Allan, E. J., Hoischen, C. & Gumpert, J. Bacterial L-forms. *Adv. Appl. Microbiol*. (in the press).
6. Joseleau-Petit, D. *et al.* Unstable *Escherichia coli* L-forms revisited: Growth requires peptidoglycan synthesis. *J. Bacteriol.* **189,** 6512–6520 (2007).
7. Allan, E. J. Induction and cultivation of a stable L-form of *Bacillus subtilis. J. Appl. Bacteriol.* **70,** 339–343 (1991).
8. Daniel, R. A. & Errington, J. DNA sequence of the *murE-murD* region of *Bacillus subtilis* 168. *Gen. Microbiol.* **139,** 361–370 (1993).
9. Leaver, M. & Errington, J. Roles for MreC and MreD proteins in helical growth of the cylindrical cell wall in *Bacillus subtilis. Mol. Microbiol.* **57,** 1196–1209 (2005).
10. Murray, T., Popham, T. & Setlow, P. *Bacillus subtilis* cells lacking penicillin-binding protein 1 require increased levels of divalent cations for growth. *J. Bacteriol.* **180,** 4555–4563 (1998).
11. Rogers, H. J., Thurman, P. F. & Buxton, R. S. Magnesium and anion requirements of *rodB* mutants of *Bacillus subtilis. J. Bacteriol.* **125,** 556–564 (1976).
12. Formstone, A. & Errington, J. A magnesium-dependent *mreB* null mutant: Implications for the role of *mreB* in *Bacillus subtilis. Mol. Microbiol.* **55,** 1646–1657 (2005).
13. Gilpin, R. W., Young, F. E. & Chatterjee, A. N. Characterization of a stable L-form of *Bacillus subtilis* 168. *J. Bacteriol.* **113,** 486–499 (1973).
14. Burmeister, H. R. & Hesseltine, C. W. Induction and propagation of a *Bacillus subtilis* L-form in natural and synthetic media. *J. Bacteriol.* **95,** 1857–1861 (1968).
15. Gilpin, R. W. & Nagy, S. S. Time-lapse photography of *Bacillus subtilis* L-forms replicating in liquid medium. *J. Bacteriol.* **127,** 1018–1021 (1976).
16. Margulies, M. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437,** 376–380 (2005).
17. Moszer, I. *et al.* SubtiList: The reference database for the *Bacillus subtilis* genome. *Nucleic Acids Res.* **30,** 62–65 (2002).
18. Fujisaki, S. *et al.* Cloning and nucleotide sequence of the *ispA* gene responsible for farnesyl diphosphate synthase activity in *Escherichia coli. J. Biochem.* **108,** 995–1000 (1990).
19. Hosfield, D. J. *et al.* Structural basis for bisphosphonate-mediated inhibition of isoprenoid biosynthesis. *J. Biol. Chem.* **279,** 8526–8529 (2004).
20. Joly, A. & Edwards, P. A. Effect of site-directed mutagenesis of conserved aspartate and arginine residues upon farnesyl diphosphate synthase activity. *J. Biol. Chem.* **268,** 26983–26989 (1993).
21. Domingue, G. J. Sr & Woody, H. B. Bacterial persistence and expression of disease. *Clin. Microbiol. Rev.* **10,** 320–344 (1997).
22. Wyrick, P. B. & Rogers, H. J. Isolation and characterization of cell wall-defective variants of *Bacillus subtilis* and *Bacillus licheniformis. J. Bacteriol.* **116,** 456–465 (1973).
23. Ward, J. B. Peptidoglycan synthesis in L-phase variants of *Bacillus licheniformis* and *Bacillus subtilis. J. Bacteriol.* **124,** 668–678 (1975).
24. King, J. R., Prescott, B. & Caldes, G. Lack of murein in a formamide-insoluble fraction from the stable L-form of *Streptococcus faecium. J. Bacteriol.* **102,** 196–197 (1970).
25. Beall, B. & Lutkenhaus, J. FtsZ in *Bacillus subtilis* is required for vegetative septation and for asymmetric septation during sporulation. *Genes Dev.* **5,** 447–455 (1991).
26. Claessen, D. *et al.* Control of the cell elongation-division cycle by shuttling of PBP1 protein in *Bacillus subtilis. Mol. Microbiol.* **68,** 1029–1046 (2008).
27. Razin, S., Yogev, D. & Naot, Y. Molecular biology and pathogenicity of mycoplasmas. *Microbiol. Mol. Biol. Rev.* **62,** 1094–1156 (1998).
28. Glockner, F. O. *et al.* Complete genome sequence of the marine planctomycete *Pirellula* sp. strain 1. *Proc. Natl Acad. Sci. USA* **100,** 8298–8303 (2003).
29. Kandler, O. & Konig, H. Cell wall polymers in Archaea (Archaebacteria). *Cell. Mol. Life Sci.* **54,** 305–308 (1998).
30. Margolin, W. Themes and variations in prokaryotic cell division. *FEMS Microbiol. Rev.* **24,** 531–548 (2000).
31. Hanczyc, M. M., Fujikawa, S. M. & Szostak, J. W. Experimental models of primitive cellular compartments: Encapsulation, growth, and division. *Science* **302,** 618–622 (2003).
32. Jones, L. J. F., Carballido-López, R. & Errington, J. Control of cell shape in bacteria: Helical, actin-like filaments in *Bacillus subtilis. Cell* **104,** 913–922 (2001).
33. Abhayawardhane, Y. & Stewart, G. C. *Bacillus subtilis* possesses a second determinant with extensive sequence similarity to the *Escherichia coli mreB* morphogene. *J. Bacteriol.* **177,** 765–773 (1995).
34. Soufo, H. J. & Graumann, P. L. Actin-like proteins MreB and Mbl from *Bacillus subtilis* are required for bipolar positioning of replication origins. *Curr. Biol.* **13,** 1916–1920 (2003).
35. Vicente, M., Gomez, M. J. & Ayala, J. A. Regulation of transcription of cell division genes in the *Escherichia coli dcw* cluster. *Cell. Mol. Life Sci.* **54,** 317–324 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Swan for assistance with bioinformatics; A. Danchin for providing unpublished sequence data; E. J. Allan for providing the L-form strain that sparked our interest; T. Davey and V. Thompson of the Electron Microscopy Research Service of Newcastle University for technical assistance with electron microscopy; S. Gruber for help with quantitative real time PCR; J.-W. Veening and H. Murray for supplying strains and advice; and D. Claessen, K. Gerdes, H. Murray and W. Vollmer for critical reading of the manuscript. M.L. was supported by the UK Biotechnology and Biological Sciences Research Council. P.D.-C. was supported by an EMBO Long-Term Fellowship.

Author Contributions M.L. and J.E. designed the research and wrote the manuscript. M.L. performed the experiments. P.D.-C. constructed plasmids and strains and contributed to discussions. J.M.C. performed the genome sequencing. R.A.D. constructed strains, and contributed to discussions and the design of the research. All authors commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.E. (jeff.errington@newcastle.ac.uk).

METHODS

Growth of strains. For growth in liquid media, strains were grown in nutrient broth (NB, Oxoid), or brain heart infusion/yeast extract (BHIYE: 37 g l^{-1} brain heart infusion with 5 g l^{-1} yeast extract, both from Oxoid). For growth on solid media, strains were grown on nutrient agar (NA, Oxoid) or in semi-solid media by adding 10 g l^{-1} type VII low gelling temperature agarose (Sigma) to liquid media. Where necessary, media were supplemented with 0.5% xylose, 1 mM IPTG, 10 μg ml^{-1} D-alanine or with 20 mM $MgCl_2$, 0.5 M sucrose and 20 mM maleic acid (MSM). As we were unsuccessful in storing L-form strains frozen in glycerol, they were propagated continuously on NA/MSM plates supplemented with 200 μg ml^{-1} penicillin G (PenG, Sigma), and re-streaked onto fresh plates every 3–4 days. Liquid cultures were initiated in NB/MSM supplemented with 200 μg ml^{-1} PenG, grown at 30 °C without shaking. To ensure adequate aeration in these conditions, the ratio of media volume:flask volume was below 1:10.

Information about the construction and genotypes of strains and plasmids are presented in Supplementary Tables 1 and 2, respectively.

Growth curves. L-form strains were inoculated into NB/MSM media supplemented with 200 μg ml^{-1} PenG. The culture was divided into portions to which the appropriate inducer or supplement was added. The culture was then aliquoted into a 96 well microtitre plate with a blank of media only. Growth of the culture at 30 °C was quantified by absorbance readings at 600 nm (A_{600}). Readings were taken every 2.75 h by a microtitre plate reader (FluoStar Galaxy, BMG Lab Technologies) with 5 min shaking before each reading. Growth curves were plotted as the average of 7 repeats.

Time-lapse microscopy. Successful time-lapse microscopy required a number of modifications to the protocols typically used for normal bacteria, mainly reflecting the slow growth rate of L-forms. L-forms were grown in 35 mm glass-bottom microwell dishes (MatTek) and visualized through the coverslip mounted over the hole in the bottom of the dish. The top surface of the coverslip was treated with 0.1% poly-lysine w/v in H_2O (Sigma) for 5 min then washed with NB/MSM. A 200 μl sample of a culture of L-forms grown in NB/MSM overnight at 30 °C was allowed to adhere to the surface of the coverslip for 10 min. Non-adherent cells were washed away with NB/MSM. 2 ml of NB/MSM was then added to the microwell dish to provide a reservoir of medium. The dish was then sealed to make an airtight chamber. To ensure that cells were not subjected to an osmotic shock, NB/MSM from the overnight culture, which had been cleared of cells by centrifugation, was used throughout the mounting procedure. The microwell dish was placed on the microscope stage and allowed to acclimatize to 30 °C for 30 min.

Electron microscopy. Samples were prepared by the Electron Microscopy Research Service of Newcastle University. Briefly, cells were fixed overnight in 2% glutaraldehyde in Sorenson's phosphate buffer (TAAB Laboratory Equipment), pH 7.4, then in 1% osmium tetroxide (Agar Scientific) for 1 h. Samples were then dehydrated in an acetone graded series before being impregnated with a graded series of epoxy resin (TAAB Laboratory Equipment) in acetone and finally embedded in 100% resin and set at 60 °C for 24 h. Cells were sectioned and counter stained with 2% uranyl acetate and lead citrate (Leica) before being imaged on a Philips CM100 Compustage Transmission Electron Microscope (FEI) with a AMT CCD camera (Deben).

Quantitative real time PCR. DNA was extracted from an exponentially growing culture of strain Bs161 (*ftsAZ*::*spc* pLOSS::*ftsAZ*) and a stationary phase culture of L-forms derived from the same strain, both were grown without selection for the plasmid. Primer pairs GATCAATCGGGGAAAGTGTG, GTAGGGCCTGTGGATTTGTG and TCCATATCCTCGCTCCTACG, ATTCTGCTGATGTGCAATGG were used to amplify the origin and terminus regions respectively, while CATTAACAGTCGGCGTTGTG, ATCCACCGCTTCTTTCATTG were used for *ftsZ*. PCR reactions were performed using Power SYBR Green PCR Master Mix (Applied Biosystems) and quantitative real time PCR was performed in a LightCycler 480 Instrument (Roche), in accordance with the manufacturer's instructions.

DNA sequencing. Genomic DNA was extracted using phenol:chloroform, and further purified using the Qiagen Genomic DNA Purification kit. Sequencing was performed using a Roche FLX Genome Sequencer. Preparation of the single-stranded template DNA library and shotgun sequencing were carried out as described by the manufacturer. Individual sequencing reads were assembled into 47 contigs using GS Mapper software (Roche) with the *B. subtilis* genomes from Subtilist[17] and A. Danchin (personal communication) as a reference sequence. Polymorphisms were filtered using an algorithm written by D. Swan of the Bioinformatics Support Unit, Newcastle University.

ARTICLES

ChIP-seq accurately predicts tissue-specific activity of enhancers

Axel Visel[1]*, Matthew J. Blow[1,2]*, Zirong Li[3], Tao Zhang[2], Jennifer A. Akiyama[1], Amy Holt[1], Ingrid Plajzer-Frick[1], Malak Shoukry[1], Crystal Wright[2], Feng Chen[2], Veena Afzal[1], Bing Ren[3], Edward M. Rubin[1,2] & Len A. Pennacchio[1,2]

A major yet unresolved quest in decoding the human genome is the identification of the regulatory sequences that control the spatial and temporal expression of genes. Distant-acting transcriptional enhancers are particularly challenging to uncover because they are scattered among the vast non-coding portion of the genome. Evolutionary sequence constraint can facilitate the discovery of enhancers, but fails to predict when and where they are active *in vivo*. Here we present the results of chromatin immunoprecipitation with the enhancer-associated protein p300 followed by massively parallel sequencing, and map several thousand *in vivo* binding sites of p300 in mouse embryonic forebrain, midbrain and limb tissue. We tested 86 of these sequences in a transgenic mouse assay, which in nearly all cases demonstrated reproducible enhancer activity in the tissues that were predicted by p300 binding. Our results indicate that *in vivo* mapping of p300 binding is a highly accurate means for identifying enhancers and their associated activities, and suggest that such data sets will be useful to study the role of tissue-specific enhancers in human biology and disease on a genome-wide scale.

The initial sequencing of the human genome[1,2], complemented by effective computational and experimental strategies for mammalian gene discovery[3,4], has resulted in a virtually complete list of protein-coding sequences. In contrast, the genomic location and function of regulatory elements that orchestrate gene expression in the developing and adult body remain more obscure, hindering studies of their contribution to developmental processes and human disease. Evolutionary constraint of non-coding sequences can predict the location of enhancers in the genome[5–12], but does not reveal when and where these enhancers are active *in vivo*. Furthermore, it has been suggested that a substantial proportion of regulatory elements is not sufficiently conserved to be detectable by comparative genomic methods[13–16].

Chromatin immunoprecipitation coupled to massively parallel sequencing (ChIP-seq) has been shown to enable genome-wide mapping of protein binding and epigenetic marks[17–22]. The ChIP-seq approach is dependent on the cross-linking of proteins to specific DNA elements, followed by antibody enrichment of the protein–DNA complexes, and high-throughput sequencing of the recovered DNA fragments. In principle, ChIP-seq using an antibody specific for an enhancer-binding protein could provide a conservation-independent approach for the identification of candidate enhancer sequences.

The acetyltransferase and transcriptional coactivator p300 is a near-ubiquitously expressed component of enhancer-associated protein assemblies and is critically required for embryonic development[23–27]. In homogeneous cell preparations, p300 has been shown to be associated with enhancers[28,29], but these *in vitro* studies provided access only to subsets of enhancers that are active in a given cell type under culture conditions, providing limited insight into their *in vivo* function. In the present study, we have determined the genome-wide occupancy of p300 in forebrain, midbrain and limb tissue isolated directly from developing mouse embryos. Using a transgenic mouse reporter assay, we show that p300 binding in these embryonic tissues predicts with high accuracy not only where enhancers are located in the genome, but also in what tissues they are active *in vivo*. Depending on the tissue type, the success rate of predicting forebrain, midbrain and limb enhancers was between 5- and 16-fold increased compared to previous studies in which such enhancers were discovered by comparative genomics[10,11].

Genome-wide mapping of p300 in tissues

To generate genome-wide maps of p300 binding *in vivo*, we microdissected forebrain, midbrain and limb tissue from more than 150 embryonic day 11.5 (E11.5) mouse embryos and performed ChIP directly from these tissue samples using a p300 antibody (Fig. 1). Immunoprecipitated DNA fragments were analysed using massively parallel sequencing and the resulting 36 base pair (bp) sequence reads were aligned to the reference mouse genome[17,30].

After appropriate quality filtering, between 2.4 and 3.6 million aligned reads obtained from each of the tissue samples were used to identify regions of the genome with significant enrichment in p300-associated DNA sequences, hereafter referred to as 'peaks' owing to their appearance in genome-wide density plots[17] (Supplementary Table 1). Using an estimated false-discovery rate (FDR) threshold of <0.01, we identified 2,543, 561 and 2,105 peaks from forebrain, midbrain and limb respectively (Supplementary Tables 2–4). Most peaks were located at least 10 kb from transcript start sites (Supplementary Fig. 1). The smaller number of peaks from midbrain is probably due to variability in the efficiency of enrichment by ChIP (Supplementary Fig. 2). Re-sampling of subsets of data suggests that the major p300-binding sites from these three tissues have been discovered, whereas with increased sequencing coverage it is anticipated that further binding sites can be identified that are occupied only in smaller subsets of cells within each tissue (Supplementary Fig. 3). Although most genomic regions with *in vivo* p300 binding were identified by peaks in a single tissue, there were 386 regions at which peaks were observed in two tissues, and 21 regions at which p300 peaks were observed in all three tissues (Supplementary Fig. 4).

[1]Genomics Division, MS 84-171, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. [2]US Department of Energy Joint Genome Institute, Walnut Creek, California 94598, USA. [3]Ludwig Institute for Cancer Research, University of California San Diego (UCSD) School of Medicine, La Jolla, California 92093, USA.
*These authors contributed equally to this work.

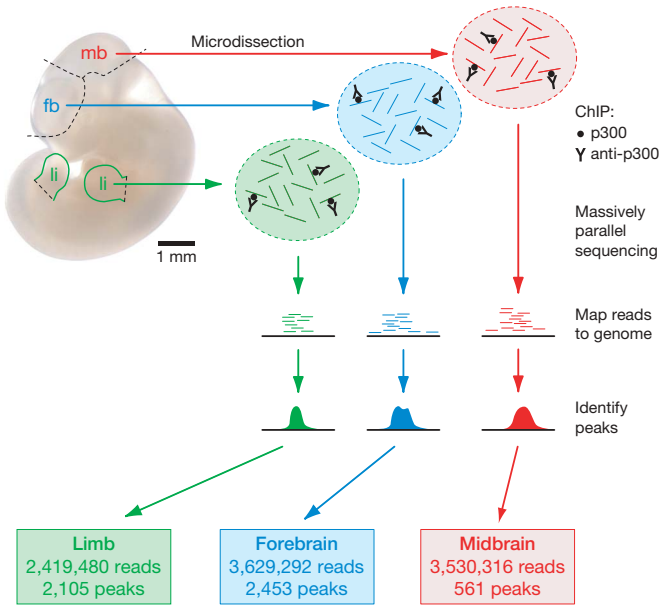


Figure 1 | Tissue dissection boundaries, overview of the ChIP-seq approach and summary of p300 results. Tissue dissection boundaries are indicated in a representative unstained E11.5 mouse embryo. For each sample, tissue was pooled from more than 150 embryos and ChIP-seq was performed with a p300-antibody. Reads obtained for each of the three tissues that unambiguously aligned to the reference mouse genome were used to define peaks (FDR < 0.01). A more comprehensive overview of sequencing and mapping results is provided in Supplementary Table 1. fb, forebrain; li, limb; mb, midbrain.

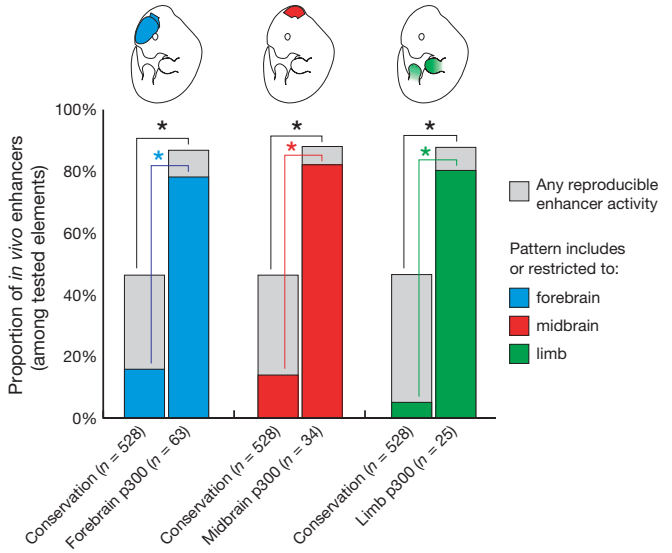


Figure 2 | p300 binding accurately predicts enhancers and their tissue-specific activity patterns. Bar height indicates the frequency of *in vivo* enhancers (reproducible at E11.5) that are active in any tissue (grey bars and coloured bars) and the fraction of enhancers in which the pattern includes or is restricted to reproducible forebrain (blue bars), midbrain (red bars) or limb activity (green bars). In each case, candidate elements predicted by p300 peaks in forebrain, midbrain or limb were compared to the frequency of the respective pattern in a background set of 528 previously tested sequences identified through extreme evolutionary conservation (combined data sets from refs 10 and 11). The component activities of elements predicted to be active in several tissues were counted separately. *$P < 0.00005$, Fisher's exact test, one-tailed.

p300 predicts enhancer activity patterns

To test directly whether p300 binding in developing mouse tissues is indicative of enhancer activity *in vivo*, we selected 86 regions with a p300 peak in at least one of the tissues for analysis in transgenic mice, comprising a total of 122 individual predictions of enhancer activity in specific tissues (Supplementary Table 5). These elements were selected blind to the identity of genes near which they are located, showed a wide range of evolutionary conservation with other vertebrate species (see Methods) and approximately reflect the genome-wide distribution properties of p300 peaks among intronic and intergenic regions, as well as their distances relative to known genes (Supplementary Fig. 1).

We cloned the human genomic sequences orthologous to these enhancer candidate regions into an enhancer reporter vector and generated transgenic mice as previously described[10,31]. For each of the 86 candidate enhancers, several independent transgenic embryos (average of $n = 8$) were assessed for reproducible reporter gene expression. A pattern was considered reproducible if the same anatomical structure was stained in three or more embryos. In almost all cases, this minimum threshold was exceeded and reproducible reporter staining in forebrain, midbrain or limb was present on average in more than 80% of the embryos obtained per construct (Supplementary Table 5).

First we determined whether p300 binding was predictive of reproducible *in vivo* enhancer activity regardless of their tissue specificity. Considering peaks from each of the three p300 data sets separately, 55 out of 63 (87%) forebrain predictions, 30 out of 34 (88%) midbrain predictions and 22 out of 25 (88%) limb predictions were active enhancers *in vivo* at E11.5 as defined by reproducible LacZ staining (Fig. 2, grey and coloured bars). Overall, 87% (75 out of 86) of the tested elements were reproducible enhancers at E11.5. This compares with a success rate for predicting enhancers of 47% (246 out of 528) from our previous studies in which elements were identified on the basis of their extreme evolutionary conservation and tested using the same transgenic mouse assay[10,11]. Thus, the rate of false-positive predictions using p300 ChIP-seq was more than fourfold lower than that with extreme evolutionary conservation (13% compared to 53% previously; $P = 4.2 \times 10^{-10}$, Fisher's exact test).

We next determined the accuracy with which p300 binding predicts the tissue in which enhancer activity will occur. Of the 63 tested elements that overlapped a forebrain p300 peak, 49 (78%) were found to have reproducible enhancer activity in the developing forebrain (Fig. 2, blue). Similarly, 28 out of 34 (82%) tested elements identified by midbrain p300 enrichment (Fig. 2, red), and 20 out of 25 (80%) tested elements identified by limb p300 enrichment (Fig. 2, green), were confirmed to be active in the predicted tissue. The 86 tested elements included 32 sequences that were identified by p300 binding in more than one tissue. Of these, 27 out of 32 (84%) were active in at least one of the predicted tissues, and 22 sequences (69%) perfectly recapitulated the predicted expression patterns (Supplementary Table 6).

To assess the degree of enrichment of enhancer activities in predicted tissues, we compared the relative frequency of enhancers for each of the three tissues examined here with a background set of 528 previously tested sequences predicted to be developmental enhancers on the basis of extreme sequence constraint that were not associated with a prior tissue specificity prediction[10,11]. For example, whereas forebrain enhancers account for only 16% (86 out of 528) of the tested elements identified through comparative approaches, 78% (49 out of 63) of elements predicted by forebrain p300 peaks were found to be active enhancers in the forebrain (Fig. 2). Forebrain predictions are therefore fivefold enriched in forebrain enhancers compared with enhancers identified through comparative approaches ($P < 1 \times 10^{-22}$). Similarly, we observed a sixfold enrichment of midbrain enhancers ($P < 1 \times 10^{-11}$) and a 16-fold enrichment of limb enhancers ($P < 1 \times 10^{-18}$) at midbrain and limb p300 peaks, respectively. Representative examples of enhancers identified by ChIP-seq are shown in Fig. 3 and detailed annotations and reproducibility across transgenic mice for all elements tested in this study can be found at http://enhancer.lbl.gov (ref. 32). Taken together, these results indicate that p300 peaks are a highly accurate predictor of *in vivo* enhancers and their spatial activity patterns.

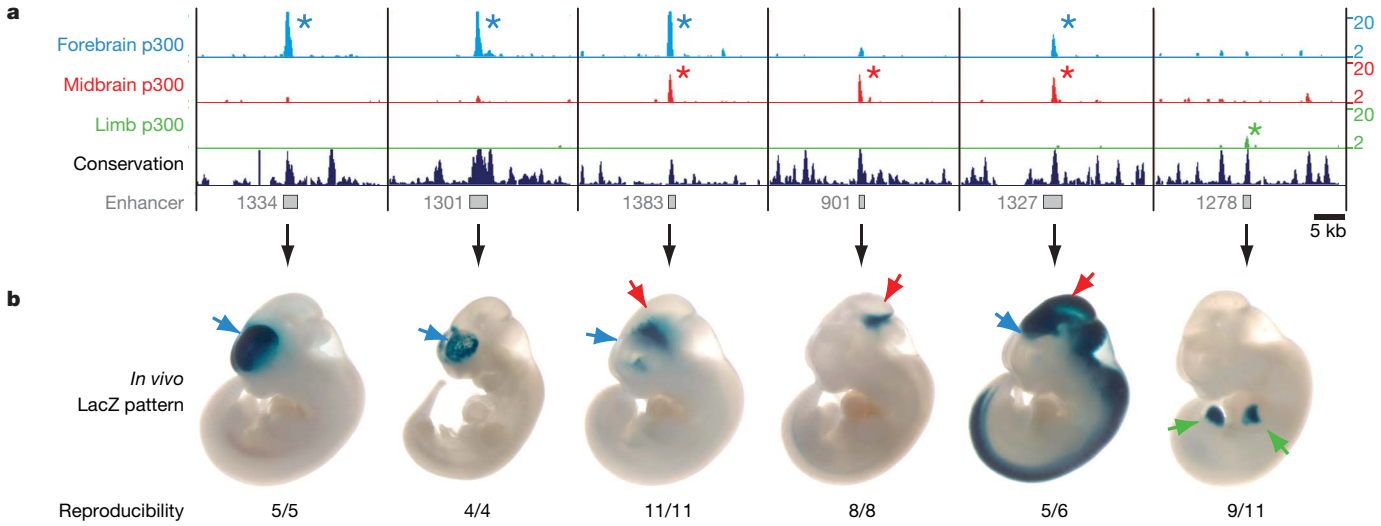


Figure 3 | Examples of successful prediction of *in vivo* enhancers by p300 binding in embryonic tissues. **a**, Coverage by extended p300 reads in forebrain (blue), midbrain (red) and limb (green). Asterisks indicate significant (FDR < 0.01) p300 enrichment in chromatin isolated from the respective tissue. Multispecies vertebrate conservation plots (black) were obtained from the UCSC genome browser[50]. Grey boxes correspond to candidate enhancer regions. Numbers at the right indicate overlapping extended reads. **b**, Representative LacZ-stained embryos with *in vivo* enhancer activity at E11.5. Reproducible staining in forebrain, midbrain and limb is indicated by arrows. Numbers show the reproducibility of LacZ reporter staining. Additional embryos obtained with each construct and genomic coordinates are available using the enhancer ID in the bottom portion of **a** at the Vista Enhancer Browser[32].

Most p300-bound regions are conserved

Previous studies have indicated a positive correlation between enhancer activity during development and non-coding sequence conservation[6,8–11,33], but it has also been suggested that not all regulatory elements in vertebrate genomes are under detectable evolutionary constraint[13–16]. To test whether p300 binding in E11.5 tissues is generally associated with evolutionarily constrained non-coding sequences, we determined if ChIP-seq reads are overall enriched at previously identified extremely conserved non-coding sequences[9,11]. We observed strong enrichment of p300 ChIP-seq reads at these conserved sequences, but not at random sites or exons (Fig. 4 and Supplementary Table 7). Vice versa, between 86% and 91% of the p300 peaks overlap sequences that are under evolutionary constraint in vertebrates[34], compared to less than 30% of size-matched random regions ($P < 1 \times 10^{-172}$, Fisher's exact test; Supplementary Fig. 5). Using a more stringent constraint threshold score, we observed that between 10% and 21% of peaks are highly constrained, compared to 1% of random regions ($P < 1 \times 10^{-82}$). These results indicate that most p300 peaks in the investigated tissues are under evolutionary constraint and support a global enrichment of p300 in highly conserved non-coding regions of the genome previously correlated with developmental enhancers.

Correlation with gene expression patterns

To examine the correlation of p300-enriched regions in embryonic tissues with the transcriptional regulation of neighbouring genes, we compared the genomic distribution of p300 peaks in E11.5 forebrain with gene expression data from this tissue. Using high-density microarrays, we identified a set of 885 genes that are overexpressed in forebrain at E11.5 compared to whole embryos (Supplementary Table 8). When we compared the genomic position of these forebrain genes to the genome-wide distribution of 2,453 forebrain-derived p300 peaks, we observed that the intervals 90 kb up- and downstream of their promoters are 2.4-fold enriched overall in p300-binding sites ($P < 0.05$, Fig. 5a). In total, 14% of all forebrain p300 peaks are located within 101 kb from a promoter of a forebrain-overexpressed gene. The most

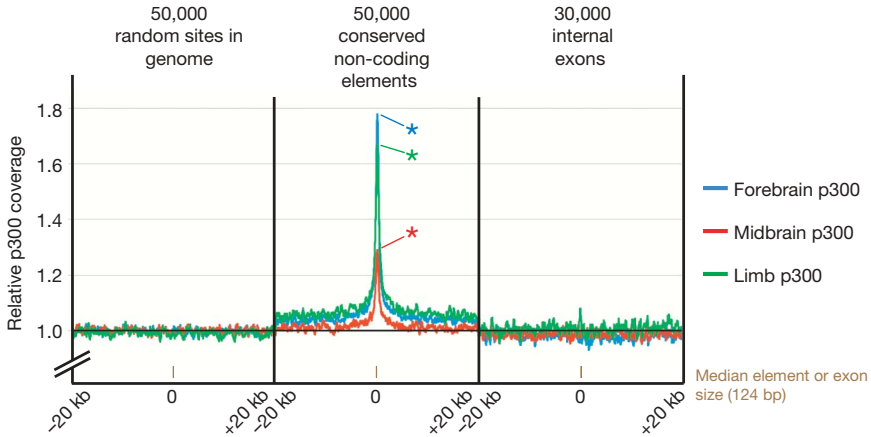


Figure 4 | In all tissues examined, p300 is enriched at highly conserved non-coding regions. We used a genome-wide set of 50,000 extremely constrained non-coding sequences identified in human-mouse-rat genome alignments[11] to assess the correlation between p300 enrichment and non-coding sequence conservation. Even though only subsets of the constrained non-coding elements are expected to be active enhancers in any given embryonic tissue, we observe strong enrichment in p300 binding in all three tissues compared to input DNA. *$P < 1 \times 10^{-100}$, Fisher's exact test. Relative p300 coverage near random sites and internal exons is shown for comparison. Brown bars indicate the median sizes of conserved elements or exons (124 bp in both cases). For further details, see Supplementary Table 7.

pronounced enrichment (4.8-fold, $P < 0.01$) was observed within 10 kb up- and downstream of promoters of forebrain-specifically expressed genes. In contrast, forebrain peaks are not enriched near genes overexpressed in other parts of the body (Fig. 5b, Supplementary Table 9). Near genes that were overexpressed fivefold or more in the forebrain, even higher enrichment of forebrain peaks was observed (11-fold enrichment within 10 kb from promoters, data not shown). We found similar enrichment of limb-derived p300 peaks near limb-overexpressed genes (Supplementary Fig. 6 and Supplementary Tables 10 and 11). These observations are consistent with the sequences bound by p300 in the forebrain or limb of day 11.5 embryos being enhancers that drive the expression of adjacent genes in these tissues at this time point.

Discussion

In the present study, we have determined the genome-wide distribution of the transcriptional coactivator protein p300 (ref. 23) using ChIP-seq[17] directly from developing mouse tissues. Notably, enrichment of p300 in different mouse tissues correctly predicted the spatial enhancer activities of human non-coding sequences in 80% of cases tested in a transgenic mouse assay, whereas absence of p300 enrichment correlated in 93% of cases with absence of enhancer activity in the respective tissue (Supplementary Table 5). The few elements that did not drive reporter gene expression in the tissue predicted by p300 ChIP-seq may represent cases in which the function of regulatory elements has diverged between the mouse sequences identified by ChIP-seq and the human orthologous regions tested in the transgenic mouse assay. In support of this hypothesis, we observed several cases in which the non-coding p300-bound region from mouse, but not the orthologous human sequence, had reproducible enhancer activities as predicted by p300 ChIP-seq from mouse tissues (data not shown). Taken together, the present approach provides a markedly improved specificity for locating enhancers in the human genome compared to conservation-based methods[10,11] and also predicts their *in vivo* activity patterns with higher accuracy than motif-based computational methods available at present (for example, refs 35, 36).

Most p300-binding regions identified in developing mouse tissues are under detectable evolutionary constraint. They typically overlap conserved non-coding sequences whose length (median of 113 bp) far exceeds that of an individual transcription factor binding site, suggesting the presence of larger functional modules. In cell culture-based chromatin studies, a sizeable fraction of non-coding regions in the human genome was found to be functional yet not constrained[13,14]. This apparent discrepancy might be due to differences in evolutionary constraint between enhancers active in developing tissues compared to those in individual cell types, but highlights the intrinsic challenge of inferring *in vivo* functionality from studies in cell culture.

A generalized picture of the epigenetic marks and proteins associated with different types of functional non-coding elements has started to emerge from genome-wide chromatin studies[13,18,28,37–41]. We can now begin to use these signatures to unravel gene regulation on a genomic scale in the context of living organisms. The highly specific approach for identification of developmental enhancers and their activity patterns presented here represents a step in this direction. Complementary *in-vivo*-derived genomic data sets may be produced in the future, covering additional embryonic stages, anatomical regions and subregions, and perhaps considering extra molecular markers[28,42–45]. Focused experiments informed by such insights will expedite studies of the genome-wide activity dynamics of enhancers in developmental, physiological and pathological processes.

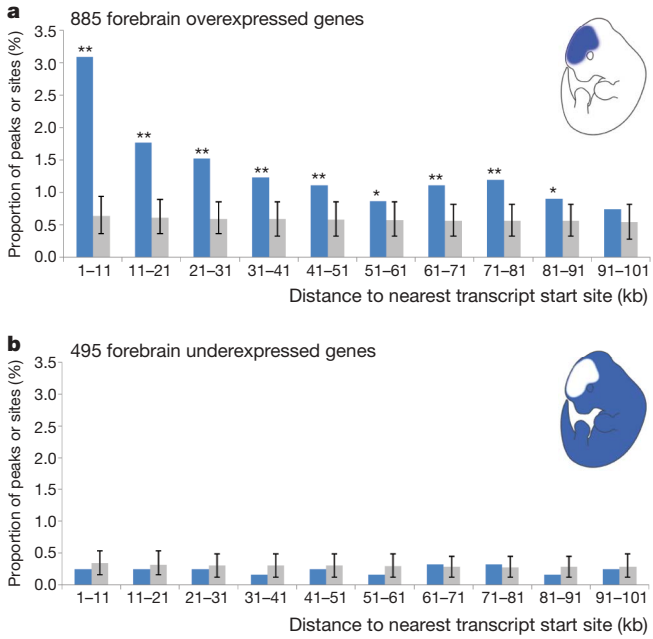


Figure 5 | p300 peaks are enriched near genes that are expressed in the same tissue. We compared the genome-wide distribution of p300-enriched regions in forebrain tissue at E11.5 with microarray expression data for forebrain at the same stage. Eight-hundred-and-eighty-five genes were forebrain-specifically overexpressed, and 495 genes were underexpressed relative to whole embryo RNA at the selected thresholds. Promoters (defined as 1 kb upstream and downstream of transcription start sites) were excluded from the analysis. Blue bars denote comparison to 2,435 forebrain-derived peaks, grey bars denote comparison to 2,435 random sites. **a**, Ten-kilobase bins up to 91 kb away from forebrain-overexpressed genes were significantly enriched in forebrain p300 peaks. **b**, No peak enrichment was observed for forebrain-underexpressed genes. Error bars indicate the 90% confidence interval on the basis of 1,000 iterations of randomized distribution. $*P < 0.05$, $**P < 0.01$, both one-tailed.

METHODS SUMMARY

Embryonic forebrain, midbrain and limb tissue was isolated from mouse embryos at E11.5. Cross-linking, chromatin isolation, sonication and immunoprecipitation using an anti-p300 antibody were performed as previously described[40,46]. ChIP DNA was further sheared by sonication, end-repaired, ligated to sequencing adapters and amplified by emulsion PCR as previously described[47]. Gel-purified amplified ChIP DNA between 300 and 500 bp was sequenced on the Illumina Genome Analyzer II platform to generate 36-bp reads.

Sequence reads were aligned to the mouse reference genome (mm9) using BLAT[48]. Uniquely aligned reads were extended to 300 bp in the 3′ direction and used to determine the read coverage at individual nucleotides at 25-bp intervals throughout the mouse genome. p300-enriched regions (peaks) with an estimated FDR of ≤ 0.01 were identified by comparison with a random distribution of the same number of reads. Candidate peaks mapping to repetitive regions were removed as probable artefacts.

Candidate regions for transgenic testing were selected based on ChIP-seq results and cover a wide spectrum of conservation. Enhancer candidate regions were amplified by PCR from human genomic DNA and cloned into an Hsp68-promoter-LacZ reporter vector as previously described[6,31]. Transgenic mouse embryos were generated and evaluated for reproducible LacZ activity at E11.5 as previously described[6].

Total RNA from E11.5 whole embryos and forebrain tissue was hybridized to GeneChip Mouse Genome 430 2.0 arrays (Affymetrix) and analysed according to the manufacturer's recommendations. Forebrain- and whole-embryo-enriched genes were identified as having at least 2.5-fold greater expression in one data set compared with the other, and a minimum signal intensity of 100. Limb-enriched genes were identified by comparison with publicly available wild-type E11.5 proximal hindlimb gene expression data (Gene Expression Omnibus (GEO) series GSE10516, samples GSM264689, GSM264690 and GSM264691)[49].

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 16 October; accepted 18 December 2008.

1. Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
2. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).

3. Burge, C. & Karlin, S. Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268,** 78–94 (1997).
4. Okazaki, Y. *et al.* Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* **420,** 563–573 (2002).
5. Marshall, H. *et al.* A conserved retinoic acid response element required for early expression of the homeobox gene *Hoxb-1*. *Nature* **370,** 567–571 (1994).
6. Nobrega, M. A., Ovcharenko, I., Afzal, V. & Rubin, E. M. Scanning human gene deserts for long-range enhancers. *Science* **302,** 413 (2003).
7. de la Calle-Mustienes, E. *et al.* A functional survey of the enhancer activity of conserved non-coding sequences from vertebrate *Iroquois* cluster gene deserts. *Genome Res.* **15,** 1061–1072 (2005).
8. Woolfe, A. *et al.* Highly conserved non-coding sequences are associated with vertebrate development. *PLoS Biol.* **3,** e7 (2005).
9. Prabhakar, S. *et al.* Close sequence comparisons are sufficient to identify human cis-regulatory elements. *Genome Res.* **16,** 855–863 (2006).
10. Pennacchio, L. A. *et al. In vivo* enhancer analysis of human conserved non-coding sequences. *Nature* **444,** 499–502 (2006).
11. Visel, A. *et al.* Ultraconservation identifies a small subset of extremely constrained developmental enhancers. *Nature Genet.* **40,** 158–160 (2008).
12. Holland, L. Z. *et al.* The amphioxus genome illuminates vertebrate origins and cephalochordate biology. *Genome Res.* **18,** 1100–1111 (2008).
13. ENCODE Project Consortium. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447,** 799–816 (2007).
14. Margulies, E. H. *et al.* Analyses of deep mammalian sequence alignments and constraint predictions for 1% of the human genome. *Genome Res.* **17,** 760–774 (2007).
15. Cooper, G. M. & Brown, C. D. Qualifying the relationship between sequence conservation and molecular function. *Genome Res.* **18,** 201–205 (2008).
16. McGaughey, D. M. *et al.* Metrics of sequence constraint overlook regulatory sequences in an exhaustive analysis at phox2b. *Genome Res.* **18,** 252–260 (2008).
17. Robertson, G. *et al.* Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nature Methods* **4,** 651–657 (2007).
18. Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448,** 553–560 (2007).
19. Robertson, A. G. *et al.* Genome-wide relationship between histone H3 lysine 4 mono- and tri-methylation and transcription factor binding. *Genome Res.* **18,** 1906–1917 (2008).
20. Cuddapah, S. *et al.* Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res.* **19,** 24–32 (2009).
21. Wederell, E. D. *et al.* Global analysis of *in vivo* Foxa2-binding sites in mouse adult liver using massively parallel sequencing. *Nucleic Acids Res.* **36,** 4549–4564 (2008).
22. Valouev, A. *et al.* Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. *Nature Methods* **5,** 829–834 (2008).
23. Eckner, R. *et al.* Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. *Genes Dev.* **8,** 869–884 (1994).
24. Eckner, R., Yao, T. P., Oldread, E. & Livingston, D. M. Interaction and functional collaboration of p300/CBP and bHLH proteins in muscle and B-cell differentiation. *Genes Dev.* **10,** 2478–2490 (1996).
25. Yao, T. P. *et al.* Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell* **93,** 361–372 (1998).
26. Merika, M., Williams, A. J., Chen, G., Collins, T. & Thanos, D. Recruitment of CBP/p300 by the IFNβ enhanceosome is required for synergistic activation of transcription. *Mol. Cell* **1,** 277–287 (1998).
27. Maston, G. A., Evans, S. K. & Green, M. R. Transcriptional regulatory elements in the human genome. *Annu. Rev. Genomics Hum. Genet.* **7,** 29–59 (2006).
28. Heintzman, N. D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genet.* **39,** 311–318 (2007).
29. Xi, H. *et al.* Identification and characterization of cell type-specific and ubiquitous chromatin regulatory structures in the human genome. *PLoS Genet.* **3,** e136 (2007).
30. Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420,** 520–562 (2002).
31. Kothary, R. *et al.* A transgene containing lacZ inserted into the dystonia locus is expressed in neural tube. *Nature* **335,** 435–437 (1988).
32. Visel, A., Minovitsky, S., Dubchak, I. & Pennacchio, L. A. VISTA Enhancer Browser—a database of tissue-specific human enhancers. *Nucleic Acids Res.* **35,** D88–D92 (2007).
33. Cheng, Y. *et al.* Transcriptional enhancement by GATA1-occupied DNA segments is strongly associated with evolutionary constraint on the binding site motif. *Genome Res.* **18,** 1896–1905 (2008).
34. Siepel, A. *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15,** 1034–1050 (2005).
35. Hallikas, O. *et al.* Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell* **124,** 47–59 (2006).
36. Pennacchio, L. A., Loots, G. G., Nobrega, M. A. & Ovcharenko, I. Predicting tissue-specific enhancers in the human genome. *Genome Res.* **17,** 201–211 (2007).
37. Kim, T. H. *et al.* A high-resolution map of active promoters in the human genome. *Nature* **436,** 876–880 (2005).
38. Boyle, A. P. *et al.* High-resolution mapping and characterization of open chromatin across the genome. *Cell* **132,** 311–322 (2008).
39. Schones, D. E. & Zhao, K. Genome-wide approaches to studying chromatin modifications. *Nature Rev. Genet.* **9,** 179–191 (2008).
40. Barrera, L. O. *et al.* Genome-wide mapping and analysis of active promoters in mouse embryonic stem cells and adult organs. *Genome Res.* **18,** 46–59 (2008).
41. Chen, X. *et al.* Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133,** 1106–1117 (2008).
42. Kwok, R. P. *et al.* Nuclear protein CBP is a coactivator for the transcription factor CREB. *Nature* **370,** 223–226 (1994).
43. Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H. & Nakatani, Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87,** 953–959 (1996).
44. Agalioti, T. *et al.* Ordered recruitment of chromatin modifying and general transcription factors to the IFN-β promoter. *Cell* **103,** 667–678 (2000).
45. Ge, K. *et al.* Transcription coactivator TRAP220 is required for PPARγ2-stimulated adipogenesis. *Nature* **417,** 563–567 (2002).
46. Li, Z. *et al.* A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells. *Proc. Natl Acad. Sci. USA* **100,** 8164–8169 (2003).
47. Blow, M. J. *et al.* Identification of the source of ancient remains through genomic sequencing. *Genome Res.* **18,** 1347–1353 (2008).
48. Kent, W. J. BLAT—the BLAST-like alignment tool. *Genome Res.* **12,** 656–664 (2002).
49. Krawchuk, D. & Kania, A. Identification of genes controlled by LMX1B in the developing mouse limb bud. *Dev. Dyn.* **237,** 1183–1192 (2008).
50. Karolchik, D. *et al.* The UCSC Genome Browser Database: 2008 update. *Nucleic Acids Res.* **36,** D773–D779 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We wish to thank R. Hosseini and S. Phouanenavong for technical support, and J. Rubenstein, J. Long, J. Choi and Y. Zhu for help with microarray experiments. This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program and by the University of California, Lawrence Berkeley National Laboratory under contract no. DE-AC02-05CH11231, Lawrence Livermore National Laboratory under contract no. DE-AC52-07NA27344, and Los Alamos National Laboratory under contract no. DE-AC02-06NA25396. L.A.P. and E.M.R. were supported by the Berkeley-PGA, under the Programs for Genomic Applications, funded by National Heart, Lung, & Blood Institute, and L.A.P. by the National Human Genome Research Institute. A.V. was supported by an American Heart Association postdoctoral fellowship. B.R. was supported by grants from the National Human Genome Research Institute and the Ludwig Institute for Cancer Research.

Author Information Reprints and permissions information is available at www.nature.com/reprints. All ChIP-seq data sets described in this study have been deposited at the National Center for Biotechnology Information (NCBI) in the Gene Expression Omnibus (GEO) database under accession number GSE13845. Correspondence and requests for materials should be addressed to L.A.P. (LAPennacchio@lbl.gov).

METHODS

Tissue dissection and chromatin immunoprecipitation. Embryonic forebrain, midbrain and limb tissue was isolated from timed-pregnant CD-1 strain mouse embryos at E11.5 by microdissection in cold PBS along the anatomical boundaries indicated in Fig. 1. Tissue samples were cross-linked (1% formaldehyde, 10 mM NaCl, 100 μM EDTA, 50 μM EGTA, 5 mM HEPES, pH 8.0) for 15 min at room temperature. Cross-linking was terminated by the addition of 125 mM glycine and cells were dissociated in a glass douncer. Chromatin isolation, sonication and immunoprecipitation were performed as previously described[40,46]. In brief, 1 mg of sonicated chromatin (OD260) was incubated with 10 μg of antibody (rabbit polyclonal anti-p300 (C-20), Santa Cruz Biotechnology) coupled to IgG magnetic beads (Dynal Biotech) overnight at 4 °C. The magnetic beads were washed eight times with RIPA buffer (50 mM HEPES, pH 8.0, 1 mM EDTA, 1% NP-40, 0.7% DOC and 0.5 M LiCl, supplemented with complete protease inhibitors from Roche Applied Science), and washed once with TE buffer (10 mM Tris, pH 8.0, 1 mM EDTA). After washing, bound DNA was eluted at 65 °C in elution buffer (10 mM Tris, pH 8.0, 1 mM EDTA and 1% SDS) for 10 min and incubated at 65 °C overnight to reverse cross-links. After the reversal of cross-linking, immunoprecipitated DNA was treated sequentially with proteinase K and RNase A, and desalted using the QIAquick PCR purification kit (Qiagen).

Amplification and Illumina sequencing of ChIP DNA. ChIP DNA was quantified by Qubit assay HS kit. Approximately 0.1 ng of each ChIP DNA sample was sheared using Sonicator XL2020 (Misonix) with a microplate horn for 10 min at 55% power output and 90% amplitude. Sheared ChIP DNA extract was end-repaired using the End-It DNA End-Repair Kit (Epicentre). Illumina adapters (56 bp and 34 bp) were ligated using T4 DNA ligase (5 U μl^{-1}, Fermentas) and recovered using a MinElute Reaction Cleanup Kit (Qiagen). Linker ligated ChIP DNA was amplified by emulsion PCR for 40 cycles as previously described[47]. Amplified ChIP DNA between 300 and 500 bp was gel purified on 2% agarose and sequenced on the Illumina Genome Analyzer II according to the manufacturer's instructions except that emulsion PCR-amplified DNA containing the GA2 sequencing adaptor was applied directly to the cluster station for bridge amplification. The resulting flow-cell was sequenced for 36 cycles to generate 36-bp reads.

Processing of Illumina sequence data. Unfiltered 36-bp Illumina sequence reads were aligned to the mouse reference genome (NCBI build 37, mm9) using BLAT[48] with optional parameters (minScore = 20, minIdentity = 80, stepSize = 5). BLAT was performed in parallel on a sge-cluster. For each read, the two highest-scoring alignments were compared and reads were rejected as repetitive unless the score of the best alignment was at least two greater than that of the second best alignment. The remaining reads were further filtered to reject those with a BLAT alignment score <21, with >1 bp insertion or deletion, or with >2 unaligned bases at the start of the read. Finally, reads with identical start sites in the mouse genome were considered likely to be duplicate sequences arising as an artefact of sample amplification or sequencing, and were counted only once. The remaining reads were classed as uniquely aligned to the mouse genome.

Uniquely aligned reads were extended to 300 bp in the 3′ direction to account for the average length of size-selected p300 ChIP fragments used for sequencing. These extended read coordinates were used to determine the read coverage at individual nucleotides at 25 bp intervals throughout the mouse genome. This data was used to produce coverage plots for visualization in the UCSC genome browser.

To identify p300-enriched regions (peaks), we compared the observed frequency of coverage depths with those expected from a random distribution of the same number of reads generated computationally as described previously[17]. In brief, the probability of observing a peak with a coverage depth of at least H reads is given by a sum of Poisson probabilities as:

$$1 - \sum_{k=0}^{H-1} \frac{e^{-\lambda}\lambda^{k}}{k!}$$

in which λ is the average genome-wide coverage of extended reads given by: (read length × number of aligned reads)/alignable genome length. To estimate the alignable genome length, one million randomly selected 36-base-polymers from the mouse genome were realigned to the mouse genome using the same alignment and filtering scheme as for reads. A total of 77.3% of 36-base-polymers were uniquely mapped back to the mouse genome, resulting in an alignable genome length of 2.107 Gb.

For each sample, we determined the read coverage depth at which the observed frequency of sites with that coverage exceeded the expected frequency by a factor of 100 (FDR ≤ 0.01). Candidate peaks were identified as sites in which the coverage exceeded this threshold, and peak boundaries were extended to the nearest flanking positions at which read coverage fell below two reads. All consecutive regions of enrichment separated by regions of continuous coverage greater than two reads were merged into a single peak. Candidate peaks mapping to chr_random contigs, centromeric regions, telomeric regions, segmental duplications, satellite repeats, ribosomal RNA repeats or regions of >70% repeat sequence, and those coinciding with enriched regions in the control sample (input DNA) were removed as probable artefacts due to misalignment of heterochromatic sequences that are not at present represented in the mouse reference genome sequence. The remaining peaks represent high-confidence p300-enriched regions and putative enhancers with activity in specific tissues.

Annotation of p300 ChIP-seq read data sets with respect to nearby genes (UCSC known genes[50]), internal exons (mouse RefSeq[51] exons >30 kb from the ends of transcripts) and conserved non-coding sequences (top 50,000 constrained non-coding human-mouse-rat conserved elements identified using GUMBY with R-ratio parameter $R = 50$; refs 9, 11) was performed using Galaxy[52] and custom Perl scripts. Annotation of p300-enriched regions with respect to UCSC known genes and vertebrate phastCons elements[34] was performed using custom Perl scripts.

Transgenic mouse enhancer assay. Candidate regions for transgenic testing were selected based on ChIP-seq results. Peaks for which human orthologous regions could not be unambiguously established and those without detectable conservation in opossum[53] were excluded from transgenic testing. Thus, the tested peaks cover a wide spectrum of conservation, but are overall more constrained than all peaks identified genome-wide (median score of 457 for all peaks versus 626 for tested peaks). Enhancer candidate regions (average size of 2.4 kb) were amplified by PCR from human genomic DNA (Clontech) and cloned into an Hsp68-promoter-LacZ reporter vector upstream of an Hsp68-promoter coupled to a LacZ reporter gene as previously described[6,31]. Candidate sequences were not cloned in any particular orientation, effectively resulting in randomized insert orientation among the test constructs. Genomic coordinates of amplified regions are reported in Supplementary Table 5. Transgenic mouse embryos were generated by pronuclear injection and F_0 embryos were collected at E11.5 and stained for LacZ activity as previously described[6]. Only patterns that were observed in at least three different embryos resulting from independent transgenic integration events of the same construct were considered reproducible (see Supplementary Table 5). To account for minor variation in separating forebrain from midbrain during tissue dissection, forebrain and midbrain p300 peaks were also considered correct predictions if the reproducible *in vivo* pattern was located in the forebrain/midbrain boundary region, whereas absence of a p300 peak was only considered a false-negative prediction if the reproducible *in vivo* pattern clearly extended beyond the boundary region.

Microarrays. Tissue was isolated from timed-pregnant CD-1 strain mouse embryos at E11.5. Forebrains were further subdivided into basal telencephalon (subpallium), dorsal telencephalon (pallium), and diencephalon, which were processed separately in subsequent steps. For comparison, whole embryos (littermates) were collected. All samples were collected, processed and hybridized in duplicate. Total RNA was extracted using Trizol reagent (Invitrogen). Synthesis of complementary RNA, hybridization to GeneChip Mouse Genome 430 2.0 arrays (Affymetrix) and analysis of hybridization results was performed according to the manufacturer's recommendations. For each sample, the average expression value from duplicates was used for downstream analyses. Forebrain-enriched genes were defined as those with expression at least 2.5-fold greater expression in at least one of the three forebrain regions compared with the whole embryo, and with a minimum signal intensity of 100. Whole embryo-enriched genes are defined as those with at least 2.5-fold greater expression in the whole embryo than in each of the three forebrain regions, and a minimum signal intensity of 100. Distances between p300 peaks and the 5′ end of Affymetrix consensus complementary DNA sequences from mouse MOE430 (A and B) aligned to the mouse reference genome (mm9) were used to determine the closest forebrain-enriched and whole embryo-enriched genes (Supplementary Tables 8 and 9). The same procedure was used to analyse the correlation of limb p300 peaks with limb gene expression, except that limb expressed genes were identified by comparison of publicly available wild-type E11.5 proximal hindlimb gene expression data (GEO series GSE10516, samples GSM264689, GSM264690 and GSM264691)[49], with the whole embryo gene expression data generated in the present study (Supplementary Tables 10 and 11).

Animal work. All animal work was performed in accordance with protocols reviewed and approved by the Lawrence Berkeley National Laboratory Animal Welfare and Research Committee.

51. Pruitt, K. D., Tatusova, T. & Maglott, D. R. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **35**, D61–D65 (2007).
52. Giardine, B. *et al.* Galaxy: a platform for interactive large-scale genome analysis. *Genome Res.* **15**, 1451–1455 (2005).
53. Mikkelsen, T. S. *et al.* Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. *Nature* **447**, 167–177 (2007).

LETTERS

Tunable delay of Einstein–Podolsky–Rosen entanglement

A. M. Marino[1], R. C. Pooser[1], V. Boyer[1,2] & P. D. Lett[1]

Entangled systems display correlations that are stronger than can be obtained classically. This makes entanglement an essential resource for a number of applications, such as quantum information processing, quantum computing and quantum communications[1,2]. The ability to control the transfer of entanglement between different locations will play a key role in these quantum protocols and enable quantum networks[3]. Such a transfer requires a system that can delay quantum correlations without significant degradation, effectively acting as a short-term quantum memory. An important benchmark for such systems is the ability to delay Einstein–Podolsky–Rosen (EPR) levels of entanglement and to be able to tune the delay. EPR entanglement is the basis for a number of quantum protocols, allowing the remote inference of the properties of one system (to better than its standard quantum limit) through measurements on the other correlated system. Here we show that a four-wave mixing process based on a double-lambda scheme in hot ^{85}Rb vapour allows us to obtain an optically tunable delay for EPR entangled beams of light. A significant maximum delay, of the order of the width of the cross-correlation function, is achieved. The four-wave mixing also preserves the quantum spatial correlations of the entangled beams. We take advantage of this property to delay entangled images, making this the first step towards a quantum memory for images[4].

The implementation of a quantum protocol usually requires the transfer of entanglement between different locations of the quantum processor to be properly synchronized. This requires tunable delays at least as large as the inverse of the maximum bandwidth of the signal (in our case given by the width of the cross-correlation function), that is to say, large tunable fractional delays. In general, applications such as quantum information processing and computing will require fractional delays larger than one, while quantum communications between distant locations will require much larger fractional delays. Light provides an excellent means to achieve this goal, as its propagation velocity can be controlled by tailoring the optical response of the medium. Indeed, a positive change of the index of refraction as a function of frequency leads to a reduced group velocity over a frequency range. For instance, the narrow transparency window obtained with electromagnetically induced transparency[5], the transparent region between two absorption resonances[6], or the sharp gain feature obtained with four-wave mixing in a double-lambda system[7], all lead to a large dispersion over a small frequency range and a large reduction in group velocity.

Up to now, breakthroughs in the areas of delay and storage of quantum states have been achieved in the single photon regime, for which quantum delay lines[8] and quantum memories[9–11] have been demonstrated. Continuous variables, such as the amplitude and phase of a beam of light, offer the possibility of generating entanglement deterministically, can be measured with high efficiency, and can be efficiently mapped onto atomic ensembles. These properties are crucial for the efficient implementation of a number of quantum information protocols[2] and make deterministic continuous-variable systems an attractive alternative to single photon systems. In the continuous-variable regime, results have been limited to the delay[12] and storage[13,14] of the squeezing properties of a single-mode beam of light and more recently to a small fractional delay (the ratio of the delay to the full-width at half-maximum of the cross-correlation function between the entangled beams after the delay) of the order of 0.08 for entanglement in the form of inseparability[15]. These experiments have been based on electromagnetically induced transparency, which has limited the storage efficiencies to values of the order of 15–20% (refs 13, 14). This poses an important constraint on the use of this process as a delay mechanism or quantum memory for continuous-variable entanglement. Here we show that four-wave mixing offers an alternative that preserves continuous-variable entanglement at the EPR level over a significant fractional delay and has the possibility of acting as a multi-spatial-mode delay line, allowing us to delay entangled images.

In order to characterize the entanglement between two systems, a and b, we need to analyse joint variables. For the case of the electromagnetic field, these variables correspond to the joint quadrature operators $\hat{X}_- = (\hat{X}_a - g\hat{X}_b)/\sqrt{2}$ and $\hat{Y}_+ = (\hat{Y}_a + g\hat{Y}_b)/\sqrt{2}$, which combine the amplitude ($\hat{X}_a$,$\hat{X}_b$) and phase ($\hat{Y}_a$,$\hat{Y}_b$) quadratures of the two systems. Here, g is a scaling factor which is adjusted to optimize the different entanglement criteria. Physically, the joint quadratures correspond to the amplitude difference, $\hat{X}_-$, and phase sum, $\hat{Y}_+$, between the fields for $g = 1$. Based on the noise properties of these quadratures, different degrees of continuous-variable entanglement exist. The variance, or noise, of these operators for coherent states sets the standard quantum limit (SQL) of the quadrature fluctuations. The minimum requirement for entanglement is that the states of the two systems cannot be described independently, a property known as inseparability. This property can be quantified with the inseparability parameter $\mathcal{I} = \langle \Delta\hat{X}_-^2 \rangle_N + \langle \Delta\hat{Y}_+^2 \rangle_N$, where the N subscript indicates that the variances have been normalized to the corresponding SQL. The state of the system is inseparable when $\mathcal{I} < 2$ for some value of g (ref. 16), such that having both $\langle \Delta\hat{X}_-^2 \rangle$ and $\langle \Delta\hat{Y}_+^2 \rangle$ below the SQL, or squeezed, for some g is a sufficient condition for entanglement. A stronger degree of quantum correlation is given by EPR entanglement, which can be quantified with the conditional variances of system a given system b, $V_{X_a|X_b}$ and $V_{Y_a|Y_b}$. We can define the EPR parameter $\mathcal{E}_{ab} = V_{X_a|X_b} \cdot V_{Y_a|Y_b}$ such that $\mathcal{E}_{ab} < 1$ indicates the presence of EPR entanglement[17]. In practice, it is not necessary to measure the conditional variances directly as they can be derived from the joint quadratures, as discussed in the Supplementary Information.

The experiment uses four-wave mixing in a double-lambda configuration (Fig. 1 inset), in two separate ^{85}Rb cells, to first generate[18,19]

[1]Joint Quantum Institute, National Institute of Standards and Technology and University of Maryland, Gaithersburg, Maryland 20899, USA. [2]MUARC, School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.

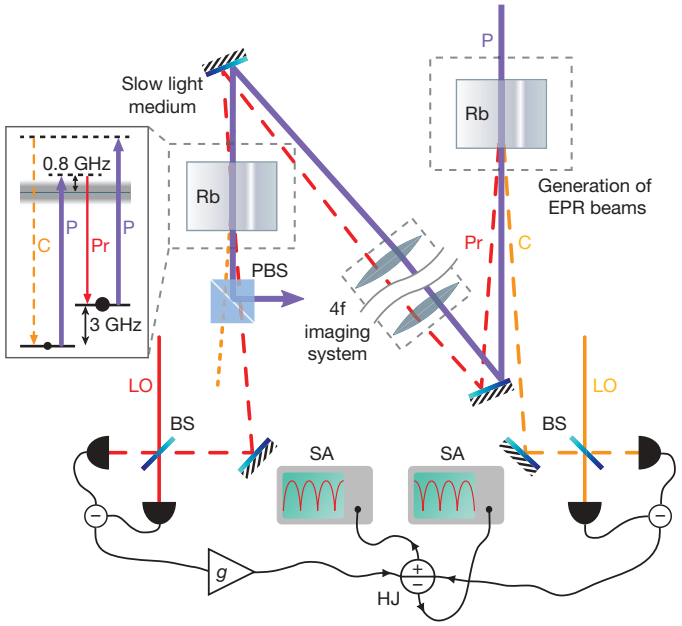


Figure 1 | Experimental set-up. In the first Rb cell (on the right), the four-wave mixing process is used to generate EPR entangled beams, which we call the probe (Pr) and the conjugate (C). The process is tuned to a gain $G \approx 4$ (at a temperature of 113 °C) and fed with vacuum states, so that vacuum twin beams are generated. After the first cell the conjugate is detected with a homodyne detector, while the probe and pump (P) are imaged with a 4f optical system into a second Rb cell. This second cell serves as the slow light medium for the probe, and is used to introduce a tunable delay between the probe and the conjugate generated by the first cell. After the second cell, the pump and the delayed probe are separated with a polarizing beam splitter (PBS) after which the probe is detected with a second homodyne detector. The sum and difference signals of the two homodyne detectors are obtained with a hybrid junction (HJ) and analysed with two spectrum analysers (SA). To obtain good spatial mode matching and phase stability when performing the homodyne detection, the local oscillators (LOs) are generated with a second four-wave mixing process in the first cell, as was done in ref. 19. The local oscillator for the probe follows a similar path as the probe through the slow light medium so that any spatial distortion is imparted to both beams. A variable electronic gain (g) is used to adjust the weighting of the homodyne detector signals in order to optimize the squeezing between the delayed probe and the conjugate. Both Rb cells have anti-reflection coated windows and are 12.5 mm long. Inset, energy level diagram for the four-wave mixing process in the D1 line of ^{85}Rb. BS, 50/50 beam splitter.

and then delay the entanglement, as shown in Fig. 1. In the first cell, entangled vacuum twin beams, called the probe and the conjugate, are generated. The second cell serves as the slow light medium for the probe, and introduces a tunable delay between the probe and the conjugate generated by the first cell. The delay can be controlled by changing the value and bandwidth of the gain, which can be done through changes in the temperature and pump power used for the four-wave mixing in the slow light cell.

Measurements of the amplitude and phase quadratures of the probe and conjugate require homodyning with local oscillators. In these measurements, the phase of the homodyne detector, or relative phase between the probe or conjugate and the corresponding local oscillator, determines the quadrature that is measured. Each beam is detected with a separate homodyne detector and the phases of the local oscillators are scanned synchronously, such that both homodyne detectors always have the same phase θ and measure the same quadrature[19], that is $\hat{X}^{\theta}_{(a,b)} = \hat{X}_{(a,b)} \cos\theta + \hat{Y}_{(a,b)} \sin\theta$. The results of both measurements are added and subtracted to obtain the joint quadratures, and the noise properties of the signals are then measured with a radio frequency spectrum analyser. The results obtained from these measurements are shown on the top row of Fig. 2 for different delays.

We use a variable electronic gain (g) in the homodyne detector for the delayed probe to adjust the weighting of the homodyne detector signals in order to optimize the squeezing measurements for the joint quadratures. This allows us to obtain the optimum value for the inseparability parameter $\mathcal{I}$. The conditional variances, required to evaluate the EPR parameter, for the conjugate given the probe can be obtained using g to optimize the squeezing with respect to the SQL of the conjugate alone[20]. In general, the values of g that minimize $\mathcal{I}$ and $\mathcal{E}_{ab}$ are not the same[21]. We can, however, use the data taken with g optimized for the inseparability parameter together with noise measurements of the individual beams to calculate the conditional variances for the EPR criterion, as described in the Supplementary Information.

The delay is characterized by the time shift in the cross-correlation function between the delayed probe and the conjugate with respect to the reference cross-correlation function, which is obtained when there is no four-wave mixing in the second cell. To calculate the correlation function of the entangled vacuum twin beams directly, we record the time traces of the photocurrents obtained from each of the homodyne detectors after filtering out the d.c. part. The correlation function has an oscillatory behaviour due to an offset of the local oscillator frequency with respect to the centre of the gain profile of the four-wave mixing process. We obtain the envelope of the correlation functions by taking data for different quadratures of the vacuum twin beams. Since the phases of the homodyne detectors are not actively stabilized, we acquire the time traces by triggering the data acquisition system at a given noise level of the signal measured by one of the spectrum analysers, which corresponds to a specific phase θ. This makes it possible to obtain data only when the homodyne detectors are measuring a specific combination of quadratures, that is, $\hat{X}^{\theta}_{(a,b)}$ (see Supplementary Information).

With our experimental parameters, the initial state is EPR entangled, with $\mathcal{E}_{ab} = 0.47(4) < 1$ and $\mathcal{I} = 0.72(3) < 2$ (Fig. 2a). As the delay is increased to 22 ns, for a fractional delay of 0.44, the squeezing in both joint quadratures is reduced to 3 dB, such that EPR entanglement is still present with $\mathcal{E}_{ab} = 0.71(9) < 1$ and $\mathcal{I} = 0.97(3) < 2$ (Fig. 2b). This point is significant, as some protocols, such as e-cloning[22], entanglement swapping[23], and teleportation of arbitrary coherent states[24], require at least 3 dB of squeezing in both joint quadratures. EPR entanglement is lost at a delay of 27 ns, as shown in Fig. 3, which corresponds to a fractional delay of 0.52. We finally increase the delay to 32 ns, for a fractional delay of 0.6. At this point a small amount of entanglement is still present, but only the inseparability criterion is satisfied with $\mathcal{I} = 1.74(3) < 2$ (Fig. 2c).

As shown in Fig. 3, longer delays require a larger gain in the slow light cell. This gain, associated with the four-wave mixing process that leads to a reduced group velocity, is the main source of excess noise responsible for the degradation of the entanglement. The gain is linked to the generation of a second conjugate which is quantum correlated with the delayed probe. As the second conjugate is not measured, this is equivalent to tracing the density matrix of the whole system over that beam, which leads to a mixed state and thus to excess noise on the delayed probe. Inseparability is more robust to sources of excess noise than EPR entanglement, as can be seen from Fig. 3, as it is independent of the purity of the state[21]. A gain of more than 2, as seen in Fig. 3, is possible while still satisfying the inseparability criterion.

The optical tunability of the system is illustrated by the fact that the delay disappears (blue correlation functions in Fig. 2) when the pump input to the slow light cell is blocked. When this is done the initial degree of entanglement is substantially recovered. As the temperature of the slow light cell is increased, absorption of the probe as it propagates through the cell increases and leads to a small degradation of the entanglement. In principle it is possible to combine both the temperature and pump power tuning capabilities of the system to obtain even larger delays. However, reducing the pump power makes the bandwidth over which a delay can be obtained smaller than the squeezing bandwidth of our initial state. This leads to distortion and break-up of the cross-correlation function and makes it hard to characterize the delay with our current experimental

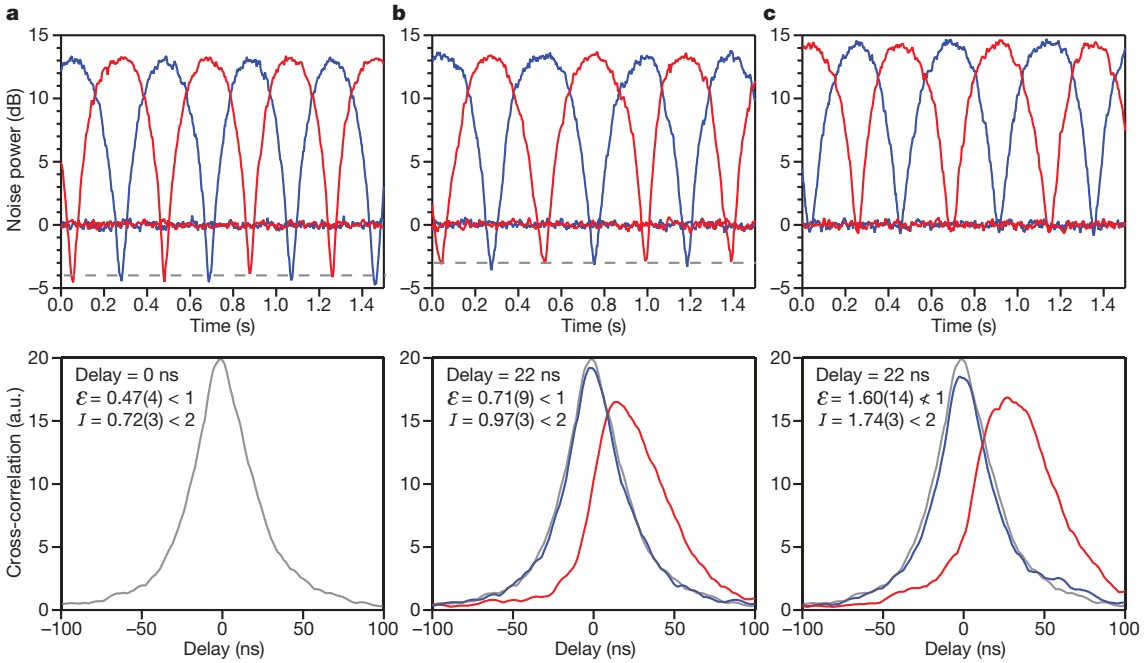


Figure 2 | Delay of EPR entanglement. Top row, the entanglement is characterized by measuring the noise power of the sum (red curves) and difference (blue curves) signals of the homodyne detectors as the phase θ is scanned. These measurements are done for a slow light cell temperature of 23 °C (**a**), 91 °C (**b**) and 100 °C (**c**). The minima of the blue and red curves give the variance of the intensity difference ($\langle\langle \Delta\hat{X}_-^2 \rangle\rangle$) and phase sum ($\langle\langle \Delta\hat{Y}_+^2 \rangle\rangle$), respectively, needed to calculate the entanglement criteria. The relative delay between the probe and conjugate is characterized by measuring the cross-correlation between their fluctuations. Bottom row, the envelopes of the cross-correlation functions for the corresponding temperatures of the slow light cell. The grey curve shows the correlation function that is obtained when there is no four-wave mixing in the slow light cell and is taken as the reference cross-correlation function for the delay measurements. This effectively corresponds to the output from the first cell. The red curves give the cross-correlation between the delayed probe and the conjugate. The delay can be controlled by changing the temperature or the pump power in the slow light cell. The blue curves show the cross-correlation functions when the pump beam for the slow light cell is blocked, which illustrates the optical tuning capabilities of the system. The envelopes of the correlation functions are calculated by taking data for 10 different phases of the homodyne detectors evenly distributed between 0 and $\pi/2$ (discussed in detail in the Supplementary Information). A data acquisition system with a rate of 1 GHz is used to record 50 time traces with 10,000 points for each trigger level, whose correlation functions are averaged. The numbers in parenthesis are combined statistical and systematic uncertainties estimated at one standard deviation. a.u., arbitrary units.

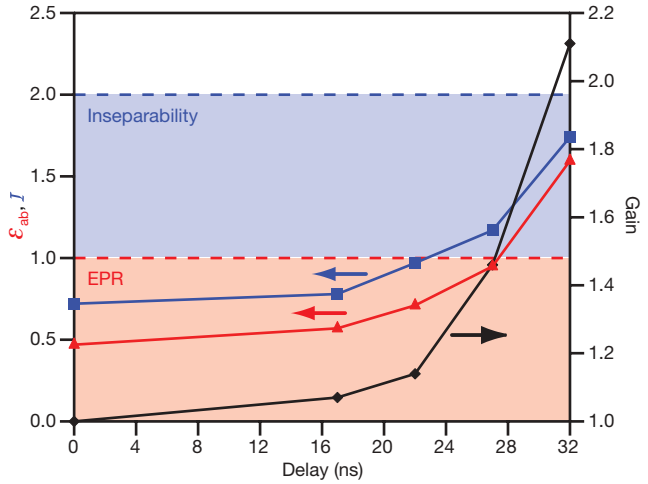


Figure 3 | Effect of the gain in the slow light cell on the quantum correlations. The black curve shows the gain in the slow light cell as a function of the delay of the probe. The red curve shows $\mathcal{E}_{ab}$ as a function of the delay. $\mathcal{E}_{ab} < 1$ (indicated by the red shaded region) indicates the presence of EPR entanglement. The blue curve shows $\mathcal{I}$ as a function of the delay. $\mathcal{I} < 2$ (indicated by the blue plus red shaded regions) indicates inseparability between the delayed probe and the conjugate. EPR entanglement is lost for a delay of about 27 ns, which corresponds to a gain of about 1.5. The inseparability is more robust to sources of excess noise, and is still present for a delay greater than 32 ns and a gain larger than 2. Longer delays require larger gains in the slow light cell, as can be seen from the black curve. This increase in gain leads to a quick deterioration of the quantum correlations.

set-up. Optical control of the delay would, however, allow for fast tuning of the propagation speed of the quantum correlated light.

The delays obtained with the four-wave mixing represent an important step towards achieving a fractional delay larger than one while maintaining the entanglement levels needed for quantum protocols in general. Even though gain degrades the quantum correlations, it should be possible to increase the fractional delay by, for example, modifying the gain profile to optimize the dispersive properties of the medium[25]. In addition, it has been recently shown that four-wave mixing in the double-lambda configuration can be used for storage in the classical regime[26]. In principle, it should be possible to use our system for the storage of quantum states.

Recently, there has been interest in extending the storage properties of quantum memories to include the transverse spatial information of the beams of light, such that the propagation of images can be controlled. Up to now, the results on this front have been limited to the delay and storage of classical images[26,27]. An important property of the four-wave mixing process used here is that it supports multiple spatial modes. This has important consequences in the generation of entangled beams of light, as it allows for subregions of the beams to be independently correlated[28] or entangled and makes it possible to obtain entanglement between complicated spatial patterns, that is, entangled images[19]. Each of the spatial modes supported by the four-wave mixing will experience gain and thus a delay. We illustrate this property by showing the delay of a particular entangled image with the shape of the '$\hbar$' symbol, as shown in Fig. 4. We have obtained a tunable relative delay between the entangled multi-spatial-mode probe and conjugate beams of up to 27 ns, for a fractional delay of 0.45, while still preserving the inseparability of the images. As a result

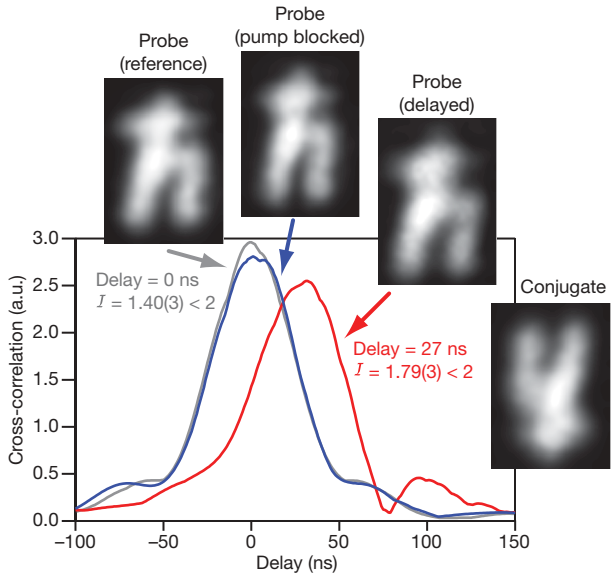


Figure 4 | Delay of an entangled image of the '$\hbar$' symbol. The four-wave mixing process preserves the spatial quantum correlations, such that for entangled images we can obtain a tunable relative delay of up to 27 ns between the probe and the conjugate. This corresponds to a fractional delay (ratio of the delay to the width of the correlation function) of 0.45. The spatial pattern for the conjugate is shown on the right side of the figure. The grey curve shows the cross-correlation between probe and conjugate when there is no delay, that is, the reference cross-correlation function. The red curve shows the cross-correlation function at the point where only a small degree of entanglement, $\mathcal{I} = 1.79(3) < 2$, is left between the delayed probe and the conjugate. The blue curve shows the cross-correlation function when the pump for the slow light cell is blocked, showing the optical tunability of the system. The four-wave mixing in the second cell introduces a slight distortion on the probe due to a cross-Kerr effect with the pump. In homodyne detection the shape of the local oscillator acts as a spatial filter that selects the spatial mode that is measured for the probe and conjugate, so that quantum correlations between arbitrary spatial modes can be characterized through the proper choice of local oscillators[19]. The images shown here correspond to the '$\hbar$' shaped local oscillators (also generated by four-wave mixing [19]) that are used to measure the entanglement of vacuum twin beams in an '$\hbar$' mode. a.u., arbitrary units.

of the phase-matching condition, given by the conservation of momentum, different spatial modes experience different gains and thus different delays. This leads to a slight broadening and distortion of the cross-correlation function.

We have demonstrated that four-wave mixing makes it possible to obtain a tunable delay of EPR entangled beams of light with a fractional delay of up to 0.52. This is a step towards applications such as distributed quantum computation, protocols based on teleportation of arbitrary quantum states, entanglement swapping, and so on. The ability to delay entangled images makes it possible to extend the intrinsic parallelism of image processing to the quantum regime. In addition, the capability of supporting multiple spatial modes offers advantages in terms of speed and reliability for quantum communication[29] and computing[30] schemes. The dependence of the delay on the pump power, combined with the spatial properties of the system, offer the possibility of using the spatial profile of the pump to independently control the delay of different spatial modes. For example, having a pump beam smaller than the probe in the cell would only introduce a delay on the overlapping regions. This would make it possible to control the transfer of a large amount of quantum resources in parallel.

Received 28 October; accepted 22 December 2008.

1. Galindo, A. & Martin-Delgado, M. A. Information and computation: Classical and quantum aspects. *Rev. Mod. Phys.* **74,** 347–423 (2002).
2. Braunstein, S. L. & van Loock, P. Quantum information with continuous variables. *Rev. Mod. Phys.* **77,** 513–577 (2005).
3. Kimble, H. J. The quantum internet. *Nature* **453,** 1023–1030 (2008).
4. Vasilyev, D. V., Sokolov, I. V. & Polzik, E. S. Quantum memory for images: A quantum hologram. *Phys. Rev. A* **77,** 020302(R) (2008).
5. Hau, L. V., Harris, S. E., Dutton, Z. & Behroozi, C. H. Light speed reduction to 17 metres per second in an ultracold atomic gas. *Nature* **397,** 594–598 (1999).
6. Camacho, R. M., Pack, M. V., Howell, J. C., Schweinsberg, A. & Boyd, R. W. Wide-bandwidth, tunable, multiple-pulse-width optical delays using slow light in cesium vapor. *Phys. Rev. Lett.* **98,** 153601 (2007).
7. Boyer, V., McCormick, C. F., Arimondo, E. & Lett, P. D. Ultraslow propagation of matched pulses by four-wave mixing in an atomic vapor. *Phys. Rev. Lett.* **99,** 143601 (2007).
8. Broadbent, C. J., Camacho, R. M., Xin, R. & Howell, J. C. Preservation of energy-time entanglement in a slow light medium. *Phys. Rev. Lett.* **100,** 133602 (2008).
9. Chaneliere, T. *et al.* Storage and retrieval of single photons transmitted between remote quantum memories. *Nature* **438,** 833–836 (2005).
10. Eisaman, M. D. *et al.* Electromagnetically induced transparency with tunable single-photon pulses. *Nature* **438,** 837–841 (2005).
11. Choi, K. S., Deng, H., Laurat, J. & Kimble, H. J. Mapping photonic entanglement into and out of a quantum memory. *Nature* **452,** 67–71 (2008).
12. Akamatsu, D. *et al.* Ultraslow propagation of squeezed vacuum pulses with electromagnetically induced transparency. *Phys. Rev. Lett.* **99,** 153602 (2007).
13. Honda, K. *et al.* Storage and retrieval of a squeezed vacuum. *Phys. Rev. Lett.* **100,** 093601 (2008).
14. Appel, J., Figueroa, E., Korystov, D., Lobino, M. & Lvovsky, A. I. Quantum memory for squeezed light. *Phys. Rev. Lett.* **100,** 093602 (2008).
15. Hétet, G. *et al.* Delay of squeezing and entanglement using electromagnetically induced transparency in a vapour cell. *Opt. Express* **16,** 7369–7381 (2008).
16. Duan, L. M., Giedke, G., Cirac, J. I. & Zoller, P. Inseparability criterion for continuous variable systems. *Phys. Rev. Lett.* **84,** 2722–2725 (2000).
17. Reid, M. D. Demonstration of the Einstein-Podolsky-Rosen paradox using nondegenerate parametric amplification. *Phys. Rev. A* **40,** 913–923 (1989).
18. McCormick, C. F., Marino, A. M., Boyer, V. & Lett, P. D. Strong low-frequency quantum correlations from a four-wave-mixing amplifier. *Phys. Rev. A* **78,** 043816 (2008).
19. Boyer, V., Marino, A. M., Pooser, R. C. & Lett, P. D. Entangled images from four-wave mixing. *Science* **321,** 544–547 (2008).
20. Ou, Z. Y., Pereira, S. F., Kimble, H. J. & Peng, K. C. Realization of the Einstein-Podolsky-Rosen paradox for continuous-variables. *Phys. Rev. Lett.* **68,** 3663–3666 (1992).
21. Bowen, W. P., Schnabel, R., Lam, P. K. & Ralph, T. C. Experimental characterization of continuous-variable entanglement. *Phys. Rev. A* **69,** 012304 (2004).
22. Weedbrook, C., Grosse, N. B., Symul, T., Lam, P. K. & Ralph, T. C. Quantum cloning of continuous-variable entangled states. *Phys. Rev. A* **77,** 052313 (2008).
23. Tan, S. M. Confirming entanglement in continuous variable quantum teleportation. *Phys. Rev. A* **60,** 2752–2758 (1999).
24. Grosshans, F. & Grangier, P. Quantum cloning and teleportation criteria for continuous quantum variables. *Phys. Rev. A* **64,** 010301(R) (2001).
25. Boyd, R. W., Gauthier, D. J., Gaeta, A. L. & Willner, A. E. Maximum time delay achievable on propagation through a slow-light medium. *Phys. Rev. A* **71,** 023801 (2005).
26. Vudyasetu, P. K., Camacho, R. M. & Howell, J. C. Storage and retrieval of multimode transverse images in hot atomic rubidium vapor. *Phys. Rev. Lett.* **100,** 123903 (2008).
27. Shuker, M., Firstenberg, O., Pugatch, R., Ron, A. & Davidson, N. Storing images in warm atomic vapor. *Phys. Rev. Lett.* **100,** 223601 (2008).
28. Boyer, V., Marino, A. M. & Lett, P. D. Generation of spatially broadband twin beams for quantum imaging. *Phys. Rev. Lett.* **100,** 143601 (2008).
29. Collins, O. A., Jenkins, S. D., Kuzmich, A. & Kennedy, T. A. B. Multiplexed memory insensitive quantum repeaters. *Phys. Rev. Lett.* **98,** 060502 (2007).
30. Tordrup, K., Negretti, A. & Molmer, K. Holographic quantum computing. *Phys. Rev. Lett.* **101,** 040501 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements R.C.P. is supported by the Intelligence Community Postdoctoral Program.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.M.M. (alberto.marino@nist.gov).

LETTERS

Ionic high-pressure form of elemental boron

Artem R. Oganov[1,2]†, Jiuhua Chen[3,4], Carlo Gatti[5], Yanzhang Ma[6], Yanming Ma[1,7], Colin W. Glass[1], Zhenxian Liu[8], Tony Yu[3], Oleksandr O. Kurakevych[9] & Vladimir L. Solozhenko[9]

Boron is an element of fascinating chemical complexity. Controversies have shrouded this element since its discovery was announced in 1808: the new 'element' turned out to be a compound containing less than 60–70% of boron, and it was not until 1909 that 99% pure boron was obtained[1]. And although we now know of at least 16 polymorphs[2], the stable phase of boron is not yet experimentally established even at ambient conditions[3]. Boron's complexities arise from frustration: situated between metals and insulators in the periodic table, boron has only three valence electrons, which would favour metallicity, but they are sufficiently localized that insulating states emerge. However, this subtle balance between metallic and insulating states is easily shifted by pressure, temperature and impurities. Here we report the results of high-pressure experiments and *ab initio* evolutionary crystal structure predictions[4,5] that explore the structural stability of boron under pressure and, strikingly, reveal a partially ionic high-pressure boron phase. This new phase is stable between 19 and 89 GPa, can be quenched to ambient conditions, and has a hitherto unknown structure (space group *Pnnm*, 28 atoms in the unit cell) consisting of icosahedral B_{12} clusters and B_2 pairs in a NaCl-type arrangement. We find that the ionicity of the phase affects its electronic bandgap, infrared adsorption and dielectric constants, and that it arises from the different electronic properties of the B_2 pairs and B_{12} clusters and the resultant charge transfer between them.

All known structures of boron contain icosahedral B_{12} clusters, with metallic-like three-centre bonds within the icosahedra and covalent two- and three-centre bonds between the icosahedra. Such bonding satisfies the octet rule and produces an insulating state, but impurity-doped boron phases are often metallic. The sensitivity of boron to impurities is evidenced by the existence of unique icosahedral boron-rich compounds such as $YB_{65.9}$, NaB_{15}, $MgAlB_{14}$, AlC_4B_{40}, NiB_{50} and PuB_{100} (refs 2, 6). In fact, probably only three of the reported boron phases correspond to the pure element[2,7,8]: rhombohedral α-B_{12} and β-B_{106} (with 12 and 106 atoms in the unit cell, respectively) and tetragonal T-192 (with 190–192 atoms per unit cell)[8]. At ambient conditions, α-B_{12} and β-B_{106} have similar static energies[9,10], but disordered β-B_{106} becomes marginally more stable (in what could seem a violation of the third law of thermodynamics) when zero-point vibrational energy is taken into account[10]. At pressures above several gigapascals, the much denser α-B_{12} phase should be more stable at all temperatures. At higher pressures, opposing effects come into play: although pressure favours metallic states and might stabilize metallic-like icosahedral clusters[11], the very low packing efficiency of atoms in icosahedral structures (34% for α-B_{12}) necessitates the destruction of the icosahedra and formation of denser phases (for example, the α-Ga-type phase[12]). In experiments, the room-temperature compression of β-B_{106} showed metastable amorphization[11] at 100 GPa and the onset of superconductivity[13] at 160 GPa. When using laser heating to overcome kinetic barriers, it was found that β-B_{106} transforms into the T-192 phase above 10 GPa at 2,280 K (ref. 14).

To further explore the intriguing high-pressure behaviour of boron, we have used 99.9999% pure β-B_{106} to synthesize (both from the melt and from the solid state) about a dozen samples containing dark-grey grains of a hitherto unknown phase of boron (see Methods for details). Single-phase samples were obtained at 12 GPa, 15 GPa and 20 GPa at temperatures of 1,800 K and 2,000 K. At 20 GPa and >2,450 K we observed the formation of the T-192 phase. Electron microprobe and X-ray microprobe analyses show that the new phase does not contain any detectable amounts of impurity atoms, and spatial homogeneity of the samples was confirmed by micro-Raman spectroscopy with a 5 μm beam. The new phase is quenchable down to ambient conditions, where X-ray diffraction patterns were collected and all the peaks indexed with an orthorhombic cell with $a = 5.0544$ Å, $b = 5.6199$ Å and $c = 6.9873$ Å. However, the structure could not be solved with our experimental diffraction data, and was found using the *ab initio* evolutionary algorithm USPEX[4,5] (see Methods, and Supplementary Information sections 1 and 2). USPEX searches for the structure with the lowest theoretical thermodynamic potential without requiring experimental information. However, the use of experimental cell parameters as constraints simplifies searches and we took advantage of this. Fully unconstrained searches were performed at higher pressures, where equilibrated experiments have not yet been performed.

We found that the structure of the new phase, which we refer to as γ-B_{28}, contains 28 atoms in the unit cell (Fig. 1, Table 1) and belongs to the space group *Pnnm*. This structure has a diffraction pattern very similar to the experimental pattern (Fig. 2). Our calculations show that it is dynamically stable (see Supplementary Information section 3), and thermodynamically more favourable that any other known or hypothetical form of boron between 19 GPa and 89 GPa at 0 K (Fig. 3). In the γ-B_{28} structure, the centres of the B_{12} icosahedra (formed by sites B2–B5; Table 1) form a slightly distorted cubic close packing as in α-B_{12}; but all octahedral voids are occupied by B_2 pairs (formed by site B1), and the new phase is thus denser than α-B_{12}. The γ-B_{28} structure resembles a NaCl-type structure, with the B_{12} icosahedra and B_2 pairs playing the roles of 'anions' and 'cations', respectively. The average intra-icosahedral bond length is 1.80 Å and the B–B bond length within the B_2 pairs is 1.73 Å.

Some metals (for example, Rb and Ba) are known to adopt under pressure broken-symmetry structures, characterized by two sublattices of the same element in different chemical roles[15]. But γ-B_{28}

[1]Laboratory of Crystallography, Department of Materials, ETH Zurich, Wolfgang-Pauli-Str. 10, CH-8093 Zurich, Switzerland. [2]Geology Department, Moscow State University, 119992 Moscow, Russia. [3]Center for the Study of Matter at Extreme Conditions and Department of Mechanical and Materials Engineering, Florida International University, Miami, Florida 33199, USA. [4]Mineral Physics Institute and Department of Geosciences, Stony Brook University, Stony Brook, New York 11794-2100, USA. [5]CNR-ISTM Istituto di Scienze e Tecnologie Molecolari, via Golgi 19, 20133 Milano, Italy. [6]Department of Mechanical Engineering, Texas University of Technology, 7th Street & Boston Avenue, Lubbock, Texas 79409, USA. [7]National Laboratory of Superhard Materials, Jilin University, Changchun 130012, China. [8]Geophysical Laboratory, Carnegie Institution of Washington, Washington DC 20015, USA. [9]LPMTM-CNRS, Université Paris Nord, Villetaneuse, F-93430, France. †Present address: Department of Geosciences and New York Center for Computational Science, Stony Brook University, Stony Brook, New York 11794-2100, USA.

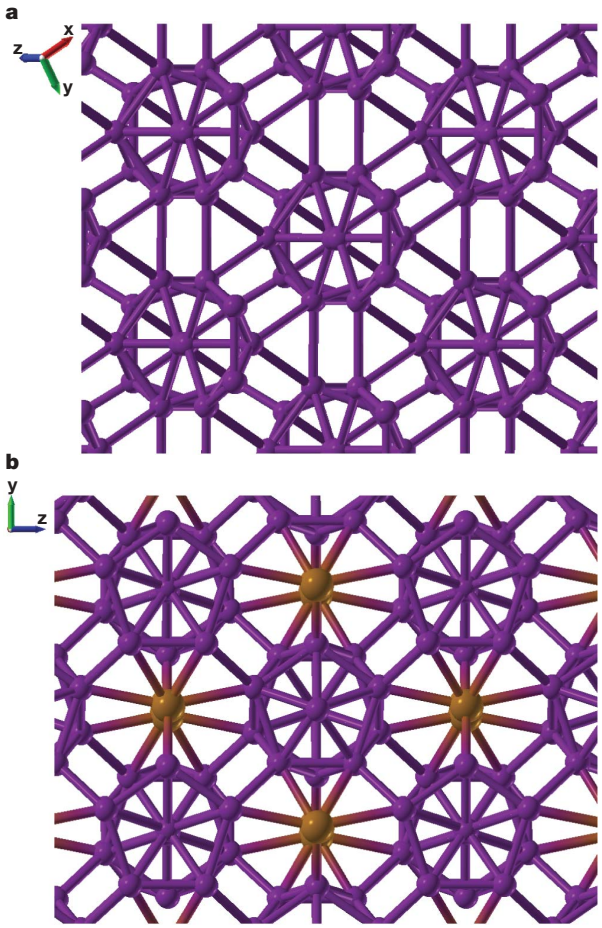


Figure 1 | Structures of α-B_{12} and γ-B_{28}. **a**, α-B_{12}; **b**, γ-B_{28}. For the partially ionic γ-B_{28} structure, two oppositely charged sublattices are marked by different colours (anionic, blue; cationic, orange). Unit cell vectors are shown.

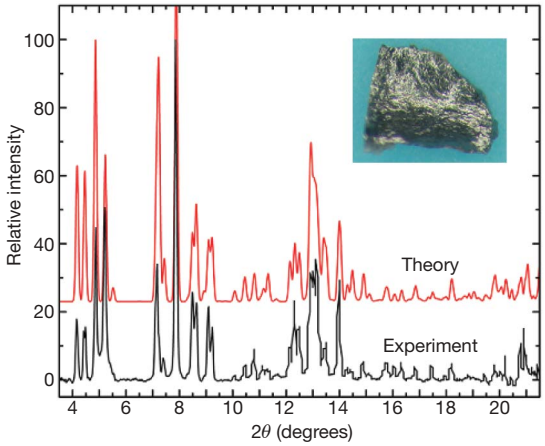


Figure 2 | Calculated and measured X-ray diffraction patterns of γ-B_{28}. X-ray wavelength is $\lambda = 0.31851$ Å. The B_2 pairs have a strong effect on the intensity of low-angle ($2\theta < 7^\circ$) peaks and are necessary for reproducing the experimental pattern. Inset, one of the samples (0.4 mm in the longest dimension).

differs in that it is non-metallic, that its two sublattices are occupied not by single atoms but by clusters (B_{12} and B_2), and charge transfer between the constituent clusters makes γ-B_{28} a boron boride, $(B_2)^{\delta+}(B_{12})^{\delta-}$. The exact magnitude of charge transfer depends on the definition of atomic charge used, but the results are all qualitatively consistent with each other. A charge estimate based on the differences in the numbers of electrons within equal atom-centred spheres (sphere radii 0.7–1.0 Å) gives $\delta \approx +0.2$, while spherically averaged Born dynamical charges give much higher values, $\delta = +2.2$. Our preferred estimates of charge transfer are based on Bader theory[16], which partitions the total electron density distribution into 'atomic' regions separated by zero-flux (that is, minimum density) surfaces and is thus physically unbiased and ensures maximum additivity and transferability of atomic properties[16]. As listed in Table 1, this approach gives $\delta = +0.34$ in local-orbital (linear combination of atomic orbitals, LCAO) and $\delta = +0.48$ in projector augmented wave (PAW) calculations. (For details of Bader analysis, see Supplementary Information section 4; for details of LCAO basis sets, see Supplementary Information sections 5 and 6.)

Evidence that these large Bader charges originate from the interaction between the B_2 and B_{12} clusters comes from the independent atom model (IAM, where the total electron density is a sum of non-interacting atomic densities), which has negligible charges that are an order of magnitude lower than when the atoms are allowed to interact. Similarly, removing the B_2 pairs from the γ-B_{28} structure while maintaining interactions, we again obtain negligible charges (within ± 0.03). The calculated electron density distribution directly shows charge transfer giving rise to a strong electron density asymmetry along the B1–B2 line with depletion near B1 and accumulation near B2 (Fig. 4a), which is typical of polar (that is, partially ionic) bonds. The bond asymmetry parameter (see Supplementary Information section 4) for this 'ionic' bond reaches 20%. Because of charge transfer, atomic volumes overall shrink (relative to the IAM) for positively charged and expand for negatively charged atoms (Table 1).

Ionicity affects many properties of γ-B_{28}: it results in large difference between high-frequency (ε_∞) and static (ε_0) dielectric constants (11.4

Table 1 | Structures of stable boron phases, optimized at 1 atm

Wyckoff position	*x*	*y*	*z*	Q_{PAW}	$Q_{LCAO\text{-}QZ}$	Q_{IAM}	$V_{LCAO\text{-}QZ}$ (a.u.)	V_{IAM} (a.u.)
γ-B_{28}*								
B1 (4g)	0.1702	0.5206	0	+0.2418	+0.1704	+0.0250	47.65	49.94
B2 (8h)	0.1606	0.2810	0.3743	−0.1680	−0.1430	−0.0153	48.46	46.25
B3 (8h)	0.3472	0.0924	0.2093	+0.0029	+0.0218	+0.0035	47.17	46.90
B4 (4g)	0.3520	0.2711	0	+0.0636	+0.0301	−0.0003	44.93	46.40
B5 (4g)	0.1644	0.0080	0	+0.0255	+0.0419	−0.0011	46.38	47.60
α-B_{12}†								
B1 (18h)	0.0103 (0.0102)	0.0103 (0.0102)	0.6540 (0.6536)	+0.0565	+0.0416	−0.0030	47.64	49.05
B2 (18h)	0.2211 (0.2212)	0.2211 (0.2212)	0.6305 (0.6306)	−0.0565	−0.0416	+0.0030	50.43	49.01
α-Ga structure‡								
B1 (8f)	0	0.1558	0.0899	0	0	0	43.08	43.08

Q, Bader charges; *V*, volumes. Values in parentheses in table and below indicate experimental data (this work, ref. 28).
* Space group *Pnnm*. $a = 5.043$ (5.054) Å, $b = 5.612$ (5.620) Å, $c = 6.921$ (6.987) Å.
† Space group $R\bar{3}$m. $a = b = c = 5.051$ (5.064) Å, $\alpha = \beta = \gamma = 58.04^\circ$ (58.10°).
‡ Space group *Cmca*. $a = 2.939$ Å, $b = 5.330$ Å, $c = 3.260$ Å.

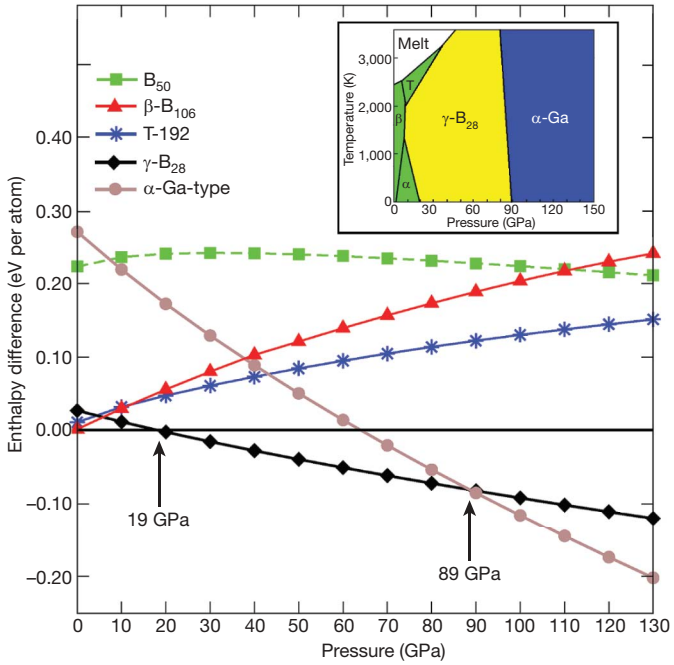


Figure 3 | Stability of boron phases. Enthalpies are shown relative to α-B_{12}. Phase transformations occur at 19 GPa (α-B_{12} to γ-B_{28}) and 89 GPa (γ-B_{28} to α-Ga-type). B_{50} is not stable for pure boron, in agreement with experiment[7], whereas γ-B_{28} has a wide stability field. For β-B_{106} we used the structural model from ref. 10. For the disordered T-192 phase, we investigated one model with full occupancy of all sites (196 atoms per cell) and three models with full occupancies of all sites but one (192 atoms per cell). The 192-atom models are energetically more favourable and accurately reproduce the observed[14] lattice parameters of T-192. Zero-point energy differences between the phases (3–7 meV per atom) are small and do not change the topology of the phase diagram. Inset, schematic phase diagram of boron, based on present results and previous experimental[14,27,29] and theoretical[10] studies. Colour-coding indicates covalent (green; α-B_{12}, β-B_{106}, T-192 phases), ionic (yellow; γ-B_{28}) and metallic (blue; α-Ga-type) solids. The phase boundary between γ-B_{28} and α-Ga-type phases is based on the static transition pressure (89 GPa) and a Clapeyron slope of -2.71 MPa K^{-1} calculated using density-functional perturbation theory and the generalized gradient approximation (-2.52 MPa K^{-1} using the local density approximation) with the ABINIT code[25]. In these calculations, the dynamical matrices were calculated on $2\times2\times2$ and $4\times2\times3$ grids in the Brillouin zone for γ-B_{28} and α-Ga-type phases, respectively, and interpolated on very dense reciprocal-space meshes. From the resulting phonon spectra we computed the entropy (S) of each phase and the high-temperature Clapeyron slope, $dP/dT = \Delta S/\Delta V$. The calculated Clapeyron slope for the α–γ transition is -4.7 MPa K^{-1}.

and 13.2, respectively), and creates the LO-TO splitting, typical of ionic crystals, and strong infrared absorption. ('LO-TO splitting' indicates the difference in frequency between the longitudinal optical and transverse optical phonons.) We measured transmission in the far-infrared (100–700 cm^{-1}) and mid-infrared (600–1,200 cm^{-1}) regions and indeed observed a number of strong absorption bands that compare well the calculated spectrum (Fig. 4b). LO-TO splitting is different for all modes and reciprocal-space directions, and the simplest parameter characterizing it is $\zeta = \sum_{i=1}^{n} \left(\omega_i^{LO}/\omega_i^{TO}\right)^2 = \varepsilon_0/\varepsilon_\infty$. In the case of non-ionic crystals, $\zeta = 1$ (we obtain 1.01 for α-B_{12}); and whenever there is charge transfer, $\zeta > 1$ (we find 1.16 for γ-B_{28} and 1.18 for GaAs). Among the individual mode splittings in γ-B_{28}, the largest one (337–375 cm^{-1}) corresponds to the most intense infrared-active mode, a nearly rigid-body opposite motion of the B_2 and B_{12} units.

Although ionic γ-B_{28} shares structural similarities with covalent α-B_{12} (Fig. 1), its electronic structure is quite different: it shows little change of the bandgap with pressure and remains even at 200 GPa an insulator with a relatively wide gap of 1.25 eV (note that density-functional

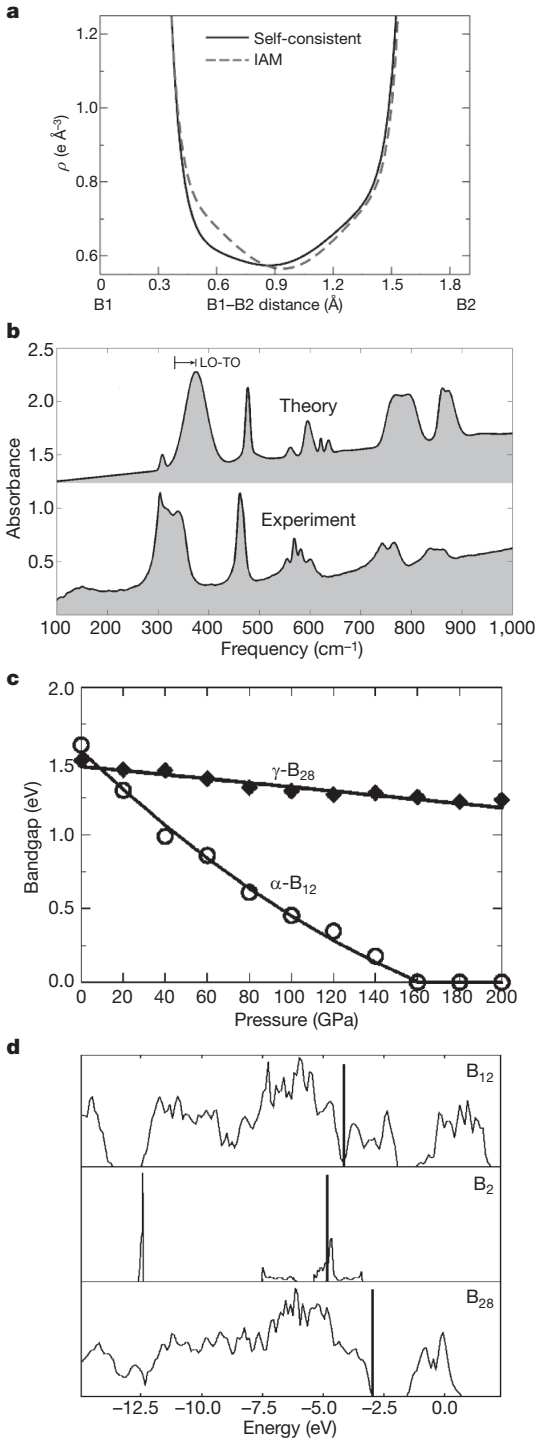


Figure 4 | Chemical bonding in γ-B_{28}. **a**, Electron density along B1-B2 line, showing polar character of the self-consistent density. IAM, independent atom model. **b**, Infrared spectra. The theoretical spectrum was obtained from mode oscillator strengths computed using density-functional perturbation theory; background was added and peaks were broadened by Gaussians with $\sigma = 2.5$–21 cm^{-1}. The agreement between theory and experiment is good, except for the intensity of the peak at 300 cm^{-1}. This mode is a sliding motion of x–y planes, and the discrepancy could be due to non-random orientation of crystallites in experimental samples, and/or surface or bulk defects (consistent with the irregular shape of that peak). Arrow denotes the largest LO-TO splitting (337–375 cm^{-1}). **c**, Pressure dependence of the bandgap of α-B_{12} and γ-B_{28}. **d**, Total electronic densities of states in γ-B_{28} and in its B_2 and B_{12} sublattices (in the same configurations as in γ-B_{28}), with Fermi levels indicated by thick vertical lines.

calculations usually underestimate bandgaps by ~40%), whereas the bandgap calculated for α-B_{12} rapidly decreases on compression and closes at ~160 GPa (Fig. 4c). These differences in behaviour arise from the presence of charge transfer in γ-B_{28} and its absence in α-B_{12}.

The origin and direction of charge transfer in γ-B_{28} is explained by the electronic properties of the B_2 and B_{12} sublattices: the latter is a p-type semiconductor (Fermi level in the valence band) and the former an n-type semiconductor (Fermi level well within the conduction band), so their interaction results in charge transfer from the n-type to the p-type sublattice and the formation of an insulating state for γ-B_{28} (Fig. 4d). Band-decomposed electron density shows that B2 atoms 'pull' the bonding orbital of $(B1)_2$ to form a polar three-centre two-electron covalent bond B2–B1–B2 (B1–B2 distances are 1.90 Å) and thereby attain a negative charge (Table 1). The top of the valence band (and bottom of the conduction band) in γ-B_{28} is clearly dominated by the B_2 pairs; this, and the spatial separation of the pairs (the nearest $B_2 \ldots B_2$ distance is 3.3 Å), explain why pressure has little effect on the bandgap of γ-B_{28}.

The example of γ-B_{28} shows that significant ionicity can occur in elemental solids, reinforcing previous theoretical suggestions (for example, hydrogen with H^+H^- molecules[17]). Ionicity appears in pure elements as a result of many-body interactions, which are strongest under pressure or when atomic orbitals are diffuse (but not diffuse enough to form metallic states)—that is, at the border between metallic and insulating states. This means that systems most favourable for exhibiting ionicity are amphoteric elements close to the Zintl line (B-Si-As-Te-At), and solids made of nanoclusters with very different electronic structures. In the case of boron studied here, our computational and experimental findings show that the ability to form B_2 and B_{12} clusters with very different electronic properties is indeed crucial for γ-B_{28} being ionic. We note that the cationic B_2^{4+} group is well known and its typical B–B distance[6] (1.70–1.75 Å in B_2F_4 and B_2Cl_4) is the same as in γ-B_{28} (1.73 Å). The B_{12} cluster is more stable as the B_{12}^{2-} anion (as in the very stable icosahedral $(B_{12}H_{12})^{2-}$ cluster[6]) because its neutral state has an unoccupied bonding orbital[18]. This orbital creates an acceptor band that is located above the valence band edge in boron-rich solids and may be partially occupied by electrons from dopant metal atoms or from other boron clusters (as detected by optical spectroscopy[19]). The B_2 pairs in γ-B_{28} thus behave as electron donors, much like the metal dopants in boron-rich borides.

γ-B_{28} remains stable up to 89 GPa, when we predict it to transform into the α-Ga-type phase (Fig. 3; see also Supplementary Fig. 2), found in our fully unconstrained variable-cell evolutionary simulations at 100 GPa and 300 GPa, in agreement with a previous intuitive proposal[12]. We note that the α-Ga-type boron phase, expected to be a poor metal exhibiting superconductivity on cooling[20], is likely to require for its synthesis experiments at >89 GPa and high temperatures (to overcome the activation barrier). The wide stability of insulating states and poor metallicity at higher pressures (at least to >300 GPa) result from the high localization of valence electrons in the boron atom (outermost orbital radius 0.78 Å; compare with 0.62 Å for C, 1.31 Å for Al, 1.25 Å for Ga).

Two centuries after the discovery of boron, we are close to understanding its phase diagram. As illustrated by the schematic phase diagram (Fig. 3 inset), ionic γ-B_{28} has a large stability field. It achieves its high density and resultant stability under pressure by using 'empty' space to host additional B_2 pairs within the B_{12} lattice. Because of the charge transfer between these two components, the new phase can be regarded as a boron boride $(B_2)^{\delta+}(B_{12})^{\delta-}$. Similar ionic interactions could also be important in other pure elemental solids (and probably also liquids) and should give rise to unusual physical properties of fundamental and potentially even practical interest.

METHODS SUMMARY

All calculations are based on density functional theory within the generalized gradient approximation[21], which gives good results for boron: for α-B_{12} at 1 atm, the predicted unit cell parameters are 0.08% smaller than experimental values (Table 1), and bond lengths are within 0.007 Å of experiment. The agreement also remains good (<1% errors in cell parameters) when zero-point vibrational pressure is taken into account.

Evolutionary structure prediction runs were done with the USPEX code[5]; the underlying structure relaxation and total-energy calculations used the frozen-core all-electron PAW method[22] as implemented in the VASP code[23]. Infrared spectra, Born charges and dielectric constants were calculated using density-functional perturbation theory[24] and plane-wave pseudopotential method, as implemented in the ABINIT code[25]. Relaxed-core all-electron LCAO calculations (using the CRYSTAL code[26]) were the main tool for analysing chemical bonding (Bader theory, projected densities of states).

Samples of γ-B_{28} were synthesized in BN capsules (at high pressures, crystalline boron does not react with BN at temperatures below 2,000 K; ref. 27) using large-volume multianvil apparatuses at Stony Brook, LPMTM-CNRS and at the Geophysical Laboratory, Carnegie Institute of Washington. Conditions of synthesis are: 12 GPa and 1,800 K (annealed for 30 min), 15 GPa and 1,800 K (annealed for 60 min), and 20 GPa and 2,000 K (annealed for 10 min). X-ray diffraction patterns of the recovered samples were collected at beamline X17C of the National Synchrotron Light Source, Brookhaven National Laboratory, at the BW5 beamline of the HASYLAB-DESY, at beamline ID27 of the ESRF, and at the LPMTM-CNRS (TEXT 3000 INEL). For additional details, see Supplementary Information section 7.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 27 January 2007; accepted 19 December 2008.
Published online 28 January 2009.

1. Petryanov-Sokolov, I. V. (ed.) *Popular Library of the Elements* Ch. 5 (Nauka, 1983).
2. Douglas, B. E. & Ho, S.-M. *Structure and Chemistry of Crystalline Solids* (Springer, 2006).
3. Chase, M. W. Jr. *J. Phys. Chem. Ref. Data* **9** (1998).
4. Oganov, A. R. & Glass, C. W. Crystal structure prediction using *ab initio* evolutionary algorithms: Principles and applications. *J. Chem. Phys.* **124,** 244704 (2006).
5. Glass, C. W., Oganov, A. R. & Hansen, N. USPEX — evolutionary crystal structure prediction. *Comput. Phys. Commun.* **175,** 713–720 (2006).
6. Wells, A. F. *Structural Inorganic Chemistry* (Clarendon, 1986).
7. Amberger, E. & Ploog, K. Bildung der Gitter des Reinen Bors. *J. Less Common Met.* **23,** 21–31 (1971).
8. Vlasse, M., Naslain, R., Kasper, J. S. & Ploog, K. Crystal structure of tetragonal boron related to α-AlB_{12}. *J. Solid State Chem.* **28,** 289–301 (1979).
9. Masago, A., Shirai, K. & Katayama-Yoshida, H. Crystal stability of α- and β-boron. *Phys. Rev. B* **73,** 104102 (2006).
10. van Setten, M. J., Uijttewaal, M. A., de Wijs, G. A. & de Groot, R. A. Thermodynamic stability of boron: The role of defects and zero point motion. *J. Am. Chem. Soc.* **129,** 2458–2465 (2007).
11. Sanz, D. N., Loubeyre, P. & Mezouar, M. Equation of state and pressure induced amorphization of β-boron from X-ray measurements up to 100 GPa. *Phys. Rev. Lett.* **89,** 245501 (2002).
12. Häussermann, U., Simak, S. I., Ahuja, R. & Johansson, B. Metal-nonmetal transition in the boron group elements. *Phys. Rev. Lett.* **90,** 065701 (2003).
13. Eremets, M. I., Struzhkin, V. W., Mao, H. K. & Hemley, R. J. Superconductivity in boron. *Science* **293,** 272–274 (2001).
14. Ma, Y. Z., Prewitt, C. T., Zou, G. T., Mao, H. K. & Hemley, R. J. High-pressure high-temperature x-ray diffraction of β-boron to 30 GPa. *Phys. Rev. B* **67,** 174116 (2003).
15. McMahon, M. I. & Nelmes, R. J. High-pressure structures and phase transformations in elemental metals. *Chem. Soc. Rev.* **35,** 943–963 (2006).
16. Bader, R. F. W. *Atoms in Molecules. A Quantum Theory* (Oxford Univ. Press, 1990).
17. Edwards, B. & Ashcroft, N. W. Spontaneous polarization in dense hydrogen. *Nature* **388,** 652–655 (1997).
18. Hayami, W. Theoretical study of the stability of AB_{12} (A = H-Ne) icosahedral clusters. *Phys. Rev. B* **60,** 1523–1526 (1999).
19. Werheit, H., Luax, M. & Kuhlmann, U. Interband and gap state related transitions in β-rhombohedral boron. *Phys. Status Solidi B* **176,** 415–432 (1993).
20. Ma, Y. M., Tse, J. S., Klug, D. D. & Ahuja, R. Electron-phonon coupling of α-Ga boron. *Phys. Rev. B* **70,** 214107 (2004).
21. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77,** 3865–3868 (1996).
22. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50,** 17953–17979 (1994).
23. Kresse, G. & Furthmüller, J. Efficiency of ab initio total-energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6,** 15–50 (1996).
24. Baroni, S., de Gironcoli, S., Dal Corso, A. & Gianozzi, P. Phonons and related crystal properties from density-functional perturbation theory. *Rev. Mod. Phys.* **73,** 515–562 (2001).
25. Gonze, X. *et al.* First-principles computation of materials properties: The ABINIT software project. *Comput. Mater. Sci.* **25,** 478–492 (2002).
26. Dovesi, R. *et al.* CRYSTAL06 User's Manual (University of Torino, 2006).

27. Solozhenko, V. L., Le Godec, Y. & Kurakevych, O. O. Solid-state synthesis of boron subnitride, B_6N: myth or reality? *C.R. Chimie* **9,** 1472–1475 (2006).
28. Will, G. & Kiefer, B. Electron deformation density in α-boron. *Z. Anorg. Allg. Chem.* **627,** 2100–2104 (2001).
29. Brazhkin, V. V., Taniguchi, T., Akaishi, M. & Popova, S. V. Fabrication of β-boron by chemical-reaction and melt-quenching methods at high pressures. *J. Mater. Res.* **19,** 1643–1648 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements A.R.O. acknowledges the Swiss National Science Foundation (grant 200021-111847/1) and the ETH Research Equipment Programme for support of this work. J.C. and T.Y. were supported by the NSF (grant EAR0711321) and the DOE (contract DE-FG02-07ER46461), Yanzhang Ma was funded by the DOE (agreement DE-FC03-03NA00144) and the NSF (grant DMR-0619215), and O.O.K. and V.L.S. were supported by the Agence Nationale de la Recherche (grant ANR-05-BLAN-0141). The use of the NSLS at Brookhaven National Laboratory was supported by the US Department of Energy under contract DE-AC02-98CH10886, and high pressure beamlines at the NSLS were supported by COMPRES under NSF cooperative agreement EAR 06-49658. Calculations were performed at CSCS (Manno), ETH Zurich, and the Joint Supercomputer Centre of the Russian Academy of Sciences.

Author Contributions A.R.O. did most of the calculations and wrote most of the paper, C.G. did most of the analysis of chemical bonding and wrote a significant part of the discussion, Yanming Ma contributed to calculations, C.W.G. wrote the first version of the USPEX code, J.C., V.L.S. and Yanzhang Ma synthesized the new phase, J.C. performed infrared absorption measurements with Z.L. and T.Y., V.L.S. did the Raman measurements and elemental analysis, V.L.S. and Yanzhang Ma did synchrotron X-ray diffraction measurements, and J.C. and O.O.K. performed the Le Bail refinement of the X-ray diffraction patterns.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.R.O. (artem.oganov@sunysb.edu).

METHODS

PAW calculations. We used a PAW potential with a [He] core (radius 1.7 atomic units, a.u.) together with a well-converged plane-wave basis set (420 eV kinetic energy cut-off) and dense **k**-point meshes of 0.1 $Å^{-1}$ resolution for sampling the Brillouin zone. Denser **k**-point meshes (resolution 0.05 $Å^{-1}$ or better) were used for the calculations of enthalpy differences (Fig. 3); electronic densities of states and bandgaps were explored with extremely dense meshes ($12 \times 12 \times 12$ for γ-B_{28} and $18 \times 18 \times 18$ for α-B_{12}). Kohn-Sham equations were solved iteratively, with a self-consistency threshold of 2×10^{-5} eV per cell. Structure relaxation proceeded until the total energy changes were below 2×10^{-4} eV per cell. We have checked that the cut-off of 420 eV gives excellent convergence for pressure (within 0.1 GPa) and energy differences (within 10^{-4} eV per atom). Evolutionary crystal structure prediction runs were performed using the PAW method, and results presented in Table 1 and Figs 1, 2, 3, 4c have been obtained using PAW calculations.

Pseudopotential calculations. Computational conditions were similar to the PAW calculations, except the use of a norm-conserving pseudopotential and the associated higher plane-wave cut-off of 884 eV to enable similar convergence. Infrared spectra (Fig. 4b), Born charges and dielectric constants were calculated using density-functional perturbation theory[24] for fully re-relaxed structures at 1 atm. Preliminary Bader analysis was also performed using pseudopotential calculations (total electron density was obtained as a sum of core and valence densities) and yielded practically the same results as the more accurate PAW and LCAO calculations reported here.

LCAO calculations. The Kohn-Sham matrix was diagonalized at an isotropic net of $8 \times 8 \times 8$ **k**-points and the same mesh was used in the reciprocal space integration for Fermi energy calculation and density matrix reconstruction. Details of the basis sets are given in Supplementary Information section 6. Results presented in Table 1 and Fig. 4a, d have been obtained using LCAO calculations. LCAO calculations were the main tool for Bader analysis reported here.

Comparison between PAW and LCAO calculations. The main differences between these calculations are in the completeness of the basis set used and in the treatment of core electrons. Well-converged plane-wave (PAW and pseudopotential) calculations use complete basis sets, but treat the core electrons as frozen. Core electrons are considered implicitly in pseudopotential calculations or explicitly in the PAW method. LCAO calculations have a somewhat less complete basis set, but treat core and valence electrons on the same footing, enabling fully variational treatment of the core orbitals (including, for example, core polarization effects). With the valence quadruple-ζ basis set used here, the consistency between the two approaches is remarkable: the energy difference between α-B_{12} and γ-B_{28} is 0.0271 eV per atom in PAW calculations, and 0.0261 eV per atom in LCAO calculations.

Experiment. Given the extreme sensitivity of boron to impurities, we used chemically pure starting material (99.9999% pure β-B_{106}), conducted experiments in BN capsules that do not react with boron at temperatures below 2,000 K (ref. 27), and checked the chemical purity and homogeneity of the recovered samples. The homogeneity of all γ-B_{28} samples was established by micro-Raman spectroscopy (Dilor XY system, 5 μm beam). X-ray electron probe microanalysis (S400, Leica/PGT Spirit and SX-50 Camebax, Cameca) of the recovered samples revealed that the impurities, if present in the γ-B_{28} samples, are at an undetectable level.

For more technical information, see the following sections of Supplementary Information: 1, details of evolutionary simulations; 2, illustration of an evolutionary simulation, making icosahedra by evolution; 3, lattice dynamics calculations and dynamical stability of the γ-B_{28} structure; 4, Bader analysis; 5, derivation of the LCAO basis sets; 6, basis set tables; 7, additional experimental details.

LETTERS

Nanomoulding with amorphous metals

Golden Kumar[1], Hong X. Tang[1] & Jan Schroers[1]

Nanoimprinting promises low-cost fabrication of micro- and nanodevices by embossing features from a hard mould onto thermoplastic materials, typically polymers with low glass transition temperature[1]. The success and proliferation of such methods critically rely on the manufacturing of robust and durable master moulds[2]. Silicon-based moulds are brittle[3] and have limited longevity[4]. Metal moulds are stronger than semiconductors, but patterning of metals on the nanometre scale is limited by their finite grain size. Amorphous metals (metallic glasses) exhibit superior mechanical properties and are intrinsically free from grain size limitations. Here we demonstrate direct nanopatterning of metallic glasses by hot embossing, generating feature sizes as small as 13 nm. After subsequently crystallizing the as-formed metallic glass mould, we show that another amorphous sample of the same alloy can be formed on the crystallized mould. In addition, metallic glass replicas can also be used as moulds for polymers or other metallic glasses with lower softening temperatures. Using this 'spawning' process, we can massively replicate patterned surfaces through direct moulding without using conventional lithography. We anticipate that our findings will catalyse the development of micro- and nanoscale metallic glass applications that capitalize on the outstanding mechanical properties, microstructural homogeneity and isotropy, and ease of thermoplastic forming exhibited by these materials[5–7].

Imprinting, embossing and moulding on the submicrometre scale are well-established techniques for manufacturing CDs (compact disks), DVDs (digital versatile disks), diffraction gratings, and polymer parts[8,9]. However, tools for imprinting on the nanometre scale (<100 nm) have only recently entered commercial production[10], and such efforts hinge on the availability of durable low-cost moulds[2] and suitable imprint resists. Candidate mould materials include essentially any type of solid material that can be precisely patterned and which exhibits sufficient strength and stability under embossing conditions. In conventional approaches, materials used for moulds and imprints exhibit separate and distinct temperature-dependent mechanical properties. The most commonly used mould materials are silicon and quartz[2,11], which retain their high strength over a wide temperature range and can be lithographically patterned. Thermoplastic polymers comprise the majority of current imprint materials. They can be easily deformed when heated above their glass transition temperature, T_g, and readily conform to a nanoscale pattern on the master mould[2].

Figure 1 compares the range of mould and imprint materials, organized by the temperature-dependent strength of each. The commonly used thermoplastic, PMMA (poly(methylmethacrylate)), decreases significantly in strength from 55 MPa at 21 °C to 0.01 MPa above 160 °C (ref. 12), becoming mouldable at low pressures under typical deformation rates used for embossing. Metals, despite their attractive mechanical properties, have not been widely used for micro- and nanoscale moulding because of difficulties associated with patterning them on a length scale smaller than their grain size. Metal patterning processes, such as micromachining, self-assembly, microcutting and etching, have not yet advanced to the level of silicon lithographic techniques. Electroforming into patterned moulds is capable of producing metal nanostructures[13]. However, electroplating of metals suffers from limited material selection and stringent process conditions and results in non-uniform mechanical properties imposed by the grain size, non-uniform deposition at sharp edges and recessed areas, surface roughness[14] and undesirable residual stresses[15].

Metallic glasses, prepared by vitrifying molten alloys, have properties that are unusual when compared to conventional metals. Bulk metallic glasses (BMGs) are a subset of glass forming alloys that can be easily vitrified and formed into relatively large (>1 mm) amorphous sections[5,6]. The absence of crystallites, grain boundaries and dislocations in the amorphous structure of bulk metallic glass results in a homogeneous and isotropic material down to the atomic scale, which displays very high strength, hardness, elastic strain limit and corrosion resistance[16]. In addition, the temperature-dependent mechanical behaviour of BMGs is unique among metals[17] (Fig. 1). Recently developed BMGs with high formability, such as $Pt_{57.5}Cu_{14.7}Ni_{5.3}P_{22.5}$ (ref. 18; Pt-BMG), $Au_{49}Ag_{5.5}Pd_{2.3}Cu_{26.9}Si_{16.3}$ (ref. 19; Au-BMG) and $Zr_{44}Ti_{11}Cu_{10}Ni_{10}Be_{25}$ (ref. 20; Zr-BMG) now offer strengths spanning

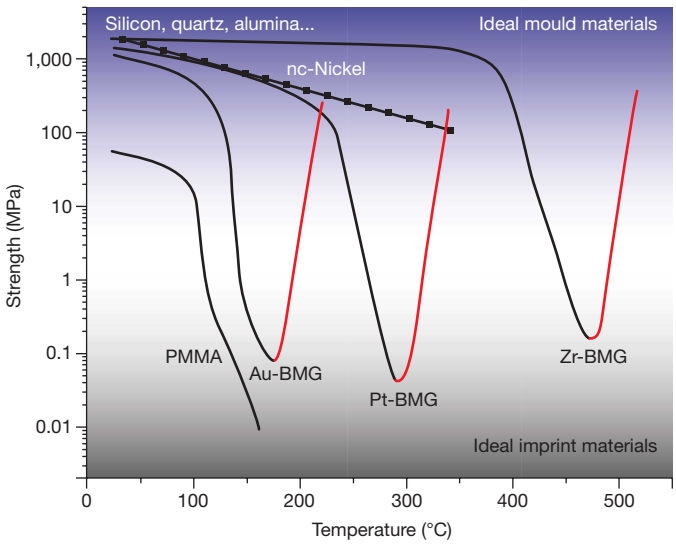


Figure 1 | Temperature-dependent strength of materials used for moulds and imprints for micro/nanoimprinting. Silicon, quartz and alumina exhibit strengths above 1,000 MPa, making them suitable for use as mould materials. The strength of nanocrystalline (nc)-nickel decreases from 1,100 MPa at room temperature to 100 MPa at 350 °C and hence limits its use to low embossing temperatures[30]. PMMA is a common imprint material due to its low forming pressure at temperatures above its T_g. The strengths of bulk metallic glasses (Au-BMG, Pt-BMG, Zr-BMG) cover a wide range from ideal mould to imprint material. They are high strength materials below their respective T_g, yet soft and mouldable in the supercooled liquid region. The strength (σ) of BMGs above T_g is calculated from their viscosity values by assuming Newtonian behaviour ($\sigma = 3\eta\dot{\varepsilon}$) and a strain rate ($\dot{\varepsilon}$) of $10^{-2}\,s^{-1}$. The strength of BMGs increases rapidly upon crystallization, as marked by the red curves.

[1]Mechanical Engineering, Yale University, New Haven, Connecticut 06511, USA.

the whole range (Fig. 1) from thermoplastics to hard moulds; some BMGs even exhibit yield strengths exceeding 5,000 MPa (ref. 21). Hence, metallic glasses can be considered high-strength materials that have the processibility of plastics in the supercooled liquid region above T_g for strain rates of the order of 10^{-2} s^{-1} (ref. 7). This is remarkably different from conventional metals, which must be heated above their melting temperatures to achieve similar processing capability. The unique softening behaviour of metallic glasses is a consequence of the crossover between the intrinsic relaxation time for flow, given by the Maxwell relaxation time, and the experimental timescale, given by the applied strain rate. The softening of metallic glasses was originally observed[22] in 1968, and first used[23] to reshape a metallic glass in 1978. Since then, this unique feature has been used for a wide range of processes, including net-shape fabrication[24,25], surface embossing[26], blow moulding[27] and writing–erasing[28]. As shown in Fig. 2, recent advances in the microforming of metallic glasses have led to a successful fabrication of parts ranging in size from several millimetres to tens of micrometres. These free standing three-dimensional components (tweezers, scalpels, a gear and a membrane)—fabricated by hot embossing the Pt-BMG on a silicon mould, followed by a subsequent hot cutting process—show the capability of metallic glass microforming to precisely replicate features with flat edges and acute angles over a wide range of length scales.

With polymer resists, the ultimate feature size is fundamentally limited by the length of the resist molecule, often about 2 nm (ref. 29). Metallic glasses are free from such internal size limiting factors, suggesting the ability to replicate atomic-level features. To date, however, moulding of metallic glasses has only demonstrated the replication of low-aspect-ratio structures (~1–3) when the lateral dimensions become smaller than 1 μm (ref. 26). The direct moulding of high-aspect-ratio structures below 100 nm has been challenging, even for BMGs with low supercooled liquid viscosities[26]. We recognized that this moulding regime is controlled by the capillary forces, not by the viscosity. The strong capillary forces can result in high moulding pressures that prevent the moulding of metallic glasses when the mould dimensions approach 100 nm or smaller. Here we have modified the Hagen–Poiseuille's law by combining the viscous and capillary contributions for an accurate description of the moulding process on nanometre length scales. This determines the required pressure to flow a supercooled liquid bulk metallic glass into a channel of diameter d and length l:

$$P = \frac{32\eta}{t}\left(\frac{l}{d}\right)^2 - \frac{4\gamma\cos\theta}{d} \quad (1)$$

Here η is the viscosity, γ is the metallic glass–vacuum interfacial energy (~1 N m^{-1}), θ is the dynamic contact angle between the supercooled liquid and the mould, t is the filling time and P is the required pressure. The second term arises from the capillary pressure, which becomes comparable to viscous pressure for $d < 1$ μm and becomes the controlling factor when d approaches 100 nm.

Figure 3a shows the pressure required to fill nanometre-sized ($d < 100$ nm) channels with aspect ratios of three for anti-wetting

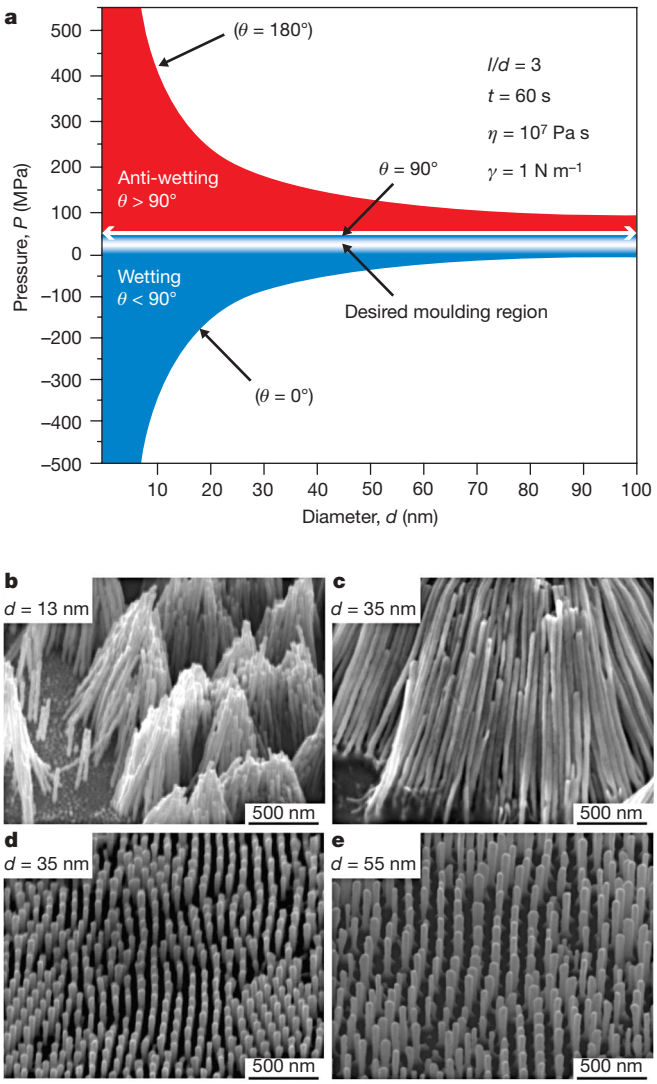


Figure 3 | Controlling the metallic glass moulding on scales smaller than 100 nm. **a**, Moulding pressure required to fill a channel of diameter d and an aspect ratio of 3 with a supercooled liquid for both anti-wetting ($\theta > 90$) and wetting ($\theta < 90$) conditions. The values used for viscosity, surface tension and filling time are 10^7 Pa s, 1 N m^{-1} and 60 s, respectively. For complete anti-wetting ($\theta = 180$), the required pressure becomes enormously high due to opposing capillary pressures. A completely wetting ($\theta = 0$) supercooled liquid spontaneously fills the mould features by capillary action. However, the partially wetting supercooled liquid permits better control over the process, and the optimal range of forming parameters is represented by the section labelled 'desired moulding region'. **b**–**e**, SEM images of Pt-BMG rods formed by embossing on porous alumina. The glassy rods of diameters 13 nm (**b**) and 35 nm (**c**) with aspect ratio exceeding 50 are formed when the Pt-BMG is heated through its supercooled liquid region (230–290 °C) under an applied pressure of 130 MPa. The process can be well controlled; glassy rods with aspect ratios of 5 fabricated by isothermal embossing at 275 °C for 60 s are depicted in **d** and **e**.

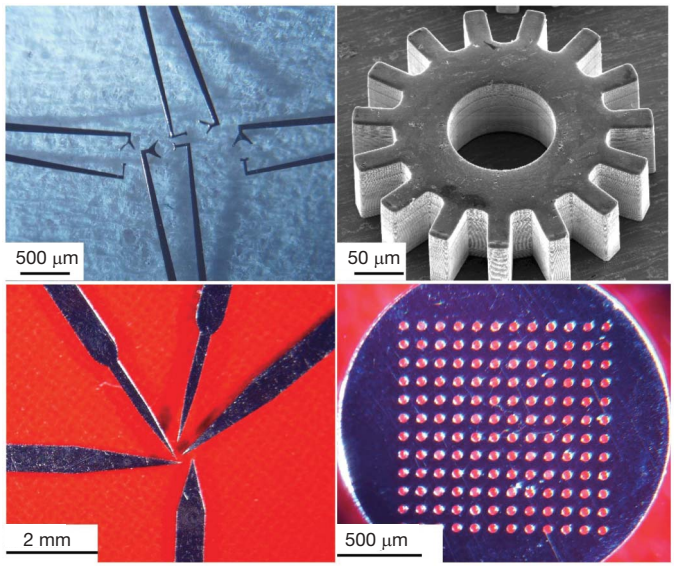


Figure 2 | Optical and scanning electron microscope images of three-dimensional microparts, including tweezers (top left), scalpels (bottom left), a gear (top right) and a membrane (bottom right). These parts were fabricated by hot embossing the Pt-BMG on an etched silicon wafer, followed by hot cutting. The precise replication of sharp edges indicates that the supercooled liquid of metallic glass readily flows into constricted areas of the mould.

($\theta > 90^\circ$) and wetting ($\theta < 90^\circ$) conditions between the supercooled liquid and the mould. For anti-wetting, the applied pressure must overcome the capillary pressure, which alone can be as high as 400 MPa for a channel of 10 nm diameter and complete anti-wetting ($\theta = 180^\circ$) conditions. In the case of complete wetting ($\theta = 0^\circ$), the supercooled liquid spontaneously fills the mould cavities, making the moulding process uncontrollable. As marked in Fig. 3a, the desired region for a controllable moulding process requires partially wetting behaviour. Therefore, the ideal properties of a metallic glass used for nanomoulding include: the ability of the supercooled liquid to partially wet the mould, access to low viscosity in the supercooled liquid region, and sluggish crystallization at the moulding temperature. Inability to meet these stringent requirements held back the development of nanomoulding of metallic glasses in the past.

Our recently developed Pt-based and Au-based metallic glasses exhibit the unique ability to partially wet typical mould materials like silicon, alumina and nickel, in addition to their low viscosities in the supercooled liquid. As we demonstrate here, this ideal combination of desirable wetting and access to low viscosity makes metallic glass nanomoulding a feasible and precisely controllable process. To demonstrate nanomoulding of metallic glasses with favourable wetting properties, the Pt-BMG was heated through its supercooled liquid region (230–290 °C) under an applied pressure of 130 MPa on porous alumina with pore diameters of 13 nm and 35 nm. Figure 3b and c shows scanning electron microscope (SEM) images of metallic glass nanorods that remained after dissolving the alumina in a KOH solution. During the drying process, the glassy rods settle into bundles, making it difficult to examine individual rods. However, the aspect ratio of each glassy rod exceeds 50, which would require an infeasible pressure of more than 4,000 MPa for an anti-wetting supercooled liquid with the same viscosity of 10^7 Pa s. It is important to understand from equation (1) that the moulding process benefits from lower viscosities only in terms of aspect ratios, whereas the smallest mouldable size is dictated by the capillary forces. The formation of high-aspect-ratio structures with sub-50-nm lateral dimensions shown in Fig. 3 are enabled by the ability of Pt-BMG to wet the mould combined with its low viscosity ($\sim 10^6$ Pa s) in the supercooled liquid state. The smallest moulded feature size is 13 nm, limited by the minimum alumina pore size, but extrapolation of our data using equation (1) suggests that features below 10 nm could be replicated with a suitable mould. The length of the moulded features can be controlled by varying the experimental parameters according to equation (1). Figure 3d and e shows SEM images of glassy nanorods with diameters of 35 nm and 55 nm fabricated by isothermal processing of Pt-based BMG on porous alumina at 275 °C for 60 s. The uniform dimensions of the moulded features suggest that nanomoulding of metallic glasses is an eminently controllable process.

The ability to replicate high-aspect-ratio nanostructures with a high strength material by direct embossing makes BMGs unique materials for nanoimprinting. Metallic glass compositions exhibit a wide variation in glass transition temperature: 128 °C for $Au_{49}Ag_{5.5}Pd_{2.3}Cu_{26.9}Si_{16.3}$ (ref. 19), 230 °C for $Pt_{57.5}Cu_{14.7}Ni_{5.3}P_{22.5}$

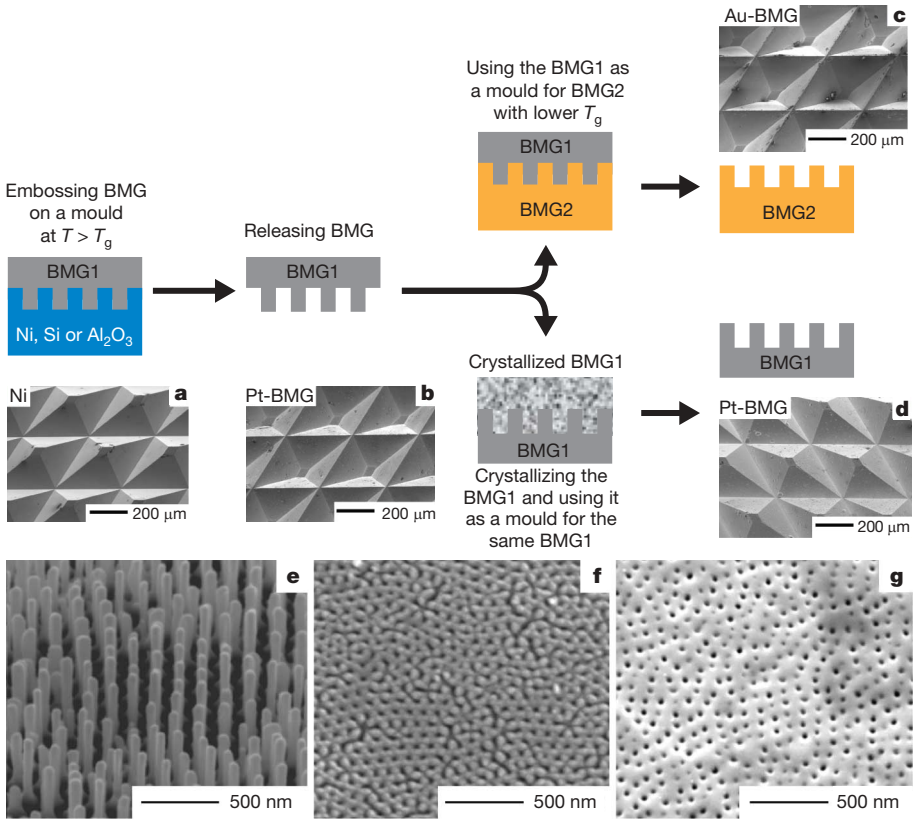


Figure 4 | Schematic and experimental illustration of a processing technique based on the unique softening behaviour of BMG. First, BMG1 is embossed on a mould fabricated by conventional techniques. The mould can be any suitable material such as Ni, Si or alumina, and an SEM image of a nickel mould used here is shown in **a** (note that only images are labelled **a**, **b** and so on). The mould and BMG1 are separated, leaving a negative pattern of the mould imprinted on BMG1. **b**, An example of pattern transfer on Pt-BMG after embossing on the Ni mould. The patterned BMG1 can be used as a mould to imprint on a lower-T_g metallic glass, BMG2. This step is demonstrated by using patterned Pt-BMG to imprint on Au-BMG (**c**). Alternatively, the crystallized BMG1 pattern can even be used as a mould for another amorphous sample of BMG1. The SEM image of a pattern transferred on Pt-BMG (**d**) from the crystallized Pt-BMG validates this concept. The process described here demonstrates the replication of 250-μm-sized features from a Ni mould into Pt-BMG and Au-BMG moulds. However, it can be extended to nanometre scale. **e**, SEM image of a Pt-BMG mould with 55-nm-diameter rods fabricated by embossing on porous alumina. **f**, Holes imprinted into PMMA using the Pt-BMG mould. **g**, The Pt-BMG mould, after crystallization, imprints holes into the same Pt-BMG, as shown by the SEM image.

(ref. 18), 350 °C for $Zr_{44}Ti_{11}Cu_{10}Ni_{10}Be_{25}$ (ref. 20), and 637 °C for $Co_{43}Fe_{20}Ta_{5.5}B_{31.5}$ (ref. 21). The broad range of softening temperatures exhibited by metallic glasses can serve as a unique toolbox for nanomoulding by employing a metallic glass with higher T_g as a mould to imprint another metallic glass with lower T_g. This sequential use of alloys with different softening temperatures can be used to replicate multiple moulds, effectively extending the original mould lifetime. The procedure is schematically illustrated along with SEM images in Fig. 4. First, a BMG (described henceforth as 'BMG1' for clarity) is hot embossed on a patterned mould (Si, Ni or alumina) prepared via lithographic or non-lithographic methods. The SEM image of the electrodeposited Ni mould used in the present example is shown in Fig. 4a. After being released from the mould, BMG1 displays inverted replicas of the mould features, shown in an SEM image of Pt-BMG separated from the Ni mould after embossing at 270 °C (Fig. 4b). Subsequently, the patterned BMG1 can be used as a mould to imprint features on polymers or another BMG (BMG2) with lower T_g. This is demonstrated by using the patterned Pt-BMG to imprint features on Au-BMG (Fig. 4c) at 160 °C. Because the strength of Pt-BMG is significantly higher than Au-BMG at 160 °C (Fig. 1), the former can be used as a mould for the latter. Alternatively, the patterned BMG1 can be crystallized and used as a mould to imprint features on a separate sample of amorphous BMG1. Figure 4d shows an SEM micrograph of the pattern transferred onto Pt-BMG using a crystallized mould of the same alloy.

The key advantage of using a crystallized mould to imprint on the same metallic glass is that the mould and the imprint are essentially of the same material with different structures, which ensures desired wetting behaviour. This allows the moulding processes to be further extended to the nanometre length scales, which otherwise were restricted to the viscosity controlled regime ($d > 100$ nm) owing to mounting capillary forces. Figure 4e shows an SEM image of Pt-BMG nanorods with 55 nm diameters fabricated by embossing on porous alumina as described earlier. The Pt-BMG substrate with nanorods was pressed into a PMMA sheet at 160 °C. Figure 4f shows the result: PMMA imprinted with 55-nm holes. Note that the image quality in Fig. 4f is limited by electron beam damage to the PMMA during SEM scanning. Finally, the Pt-BMG substrate with nanorods was crystallized and subsequently used as a mould for an amorphous sample of the same alloy. Figure 4g shows an SEM image of 55-nm-diameter holes imprinted in Pt-BMG using the crystallized Pt-BMG mould. The results summarized in Figs 3 and 4 demonstrate the ability of metallic glasses to precisely replicate mould features ranging from 250 μm to 13 nm. The two attractive features of metallic glass moulding—the ability to replicate features from tens of micrometres down to 13 nm and the possibility of using BMGs as nanotemplates—stem directly from the unique softening behaviour and homogeneous structure of metallic glasses.

The processes described here establish the capability of metallic glasses for moulding on the micrometre to nanometre scale through the use of a simple embossing technique that takes advantage of the low viscosity and favourable wetting properties of the supercooled liquid. The superior properties of metallic glasses on the nanoscale, together with the ability to form high-aspect-ratio nanostructures, suggest wide-ranging uses in different areas. Besides direct use in nanoimprinting, another potential application for nanomoulding of metallic glasses lies in the field of re-writable data storage. Our recently developed metallic glasses exhibit low viscosity in the supercooled liquid region, and thus surface features can be 'erased' under the action of surface tension alone when annealed above T_g (ref. 28). The thermal stability of the alloy is sufficient to allow thousands of write and erase cycles before the metallic glass crystallizes. Such a process also presents a unique opportunity to repair a damaged mould[28].

The nanomoulding process can be extended to metallic glass thin films, which can be directly patterned to serve as masks for photolithography. This could evade several complex steps, such as electron-beam writing and etching, involved in conventional chrome masks. Furthermore, the biocompatibility of metallic glasses and their unique ability to control structure on multiple length scales can be used to control cellular responses by fabricating medical implants with patterned surfaces. Finally, experimental access to extremely small structures should enable the study of fundamental aspects of plastic deformation in metallic glasses, such as size effects and shear transformation zones.

METHODS SUMMARY

The preparation of metallic glasses used here is described in detail in previous papers[18,19]. The parameters for thermoplastic forming experiments were selected on the basis of the crystallization and viscosity data of the metallic glasses. Three different types of moulds (Si, Ni and Al_2O_3) were used in this work. Si-based moulds were fabricated by conventional photolithographic and etching techniques. Ni-based moulds were fabricated by electroplating using the LIGA technique. Commercially available nano-porous Al_2O_3 was used as template for nanomoulding. For the moulding of very high aspect ratio nanostructures, the metallic glass was thermoplastically formed from T_g to $T_x - 10$ °C (where T_x is the onset of crystallization) under a constant load. Controlled metallic glass nanostructures for sequential nanoimprinting were generated by isothermal moulding. The metallic glasses were released from the moulds either by etching with KOH solution (for Si and Al_2O_3 moulds) or mechanical separation (for Ni moulds). The hot-cutting process was used for making free-standing three-dimensional BMG parts.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 1 June; accepted 27 November 2008.

1. Chou, S. Y., Krauss, P. R. & Renstrom, P. J. Imprint lithography with 25-nanometer resolution. *Science* **272,** 85–87 (1996).
2. Guo, L. J. Nanoimprint lithography: Methods and material requirements. *Adv. Mater.* **19,** 495–513 (2007).
3. Barbero, D. R. *et al.* High resolution nanoimprinting with a robust and reusable polymer mold. *Adv. Funct. Mater.* **17,** 2419–2425 (2007).
4. Zankovych, S., Hoffmann, T., Seekamp, J., Bruch, J. U. & Torres, C. M. S. Nanoimprint lithography: Challenges and prospects. *Nanotechnology* **12,** 91–95 (2001).
5. Greer, A. L. Metallic glasses. *Science* **267,** 1947–1953 (1995).
6. Johnson, W. L. Bulk glass-forming metallic alloys: Science and technology. *MRS Bull.* **24,** 42–56 (1999).
7. Schroers, J. & Paton, N. Amorphous metal alloys form like plastics. *Adv. Mater. Process.* **164,** 61–63 (2006).
8. Gates, B. D. *et al.* New approaches to nanofabrication: Molding, printing, and other techniques. *Chem. Rev.* **105,** 1171–1196 (2005).
9. Heckele, M. & Schomburg, W. K. Review on micro molding of thermoplastic polymers. *J. Micromech. Microeng.* **14,** R1–R14 (2004).
10. Chou, S. Y., Krauss, P. R. & Zhang, W. Sub-10 nm imprint lithography and applications. *J. Vac. Sci. Technol. B* **15,** 2897–2904 (1997).
11. Chou, S. Y., Keimel, C. & Gu, J. Ultrafast and direct imprint of nanostructures in silicon. *Nature* **417,** 835–837 (2002).
12. Palm, G., Dupaix, R. B. & Castro, J. Large strain mechanical behavior of poly(methyl methacrylate) (PMMA) near the glass transition temperature. *Trans. Am. Soc. Mech. Eng.* **128,** 559–563 (2006).
13. Hirai, Y., Harada, S. & Isaka, S. Nano-imprint lithography using replicated mold by Ni electroforming. *Jpn. J. Appl. Phys. 1* **41,** 4186–4189 (2002).
14. Haatainen, T., Majander, P., Riekkinen, T. & Ahopelto, J. Nickel stamp fabrication using step & stamp imprint lithography. *Microelectron. Eng.* **83,** 948–950 (2006).
15. Basrour, S. & Robert, L. X-ray characterization of residual stresses in electroplated nickel used in LIGA technique. *Mater. Sci. Eng. A* **288,** 270–274 (2000).
16. Schuh, C. A., Hufnagel, T. C. & Ramamurty, U. Overview No.144 – Mechanical behavior of amorphous alloys. *Acta Mater.* **55,** 4067–4109 (2007).
17. Spaepen, F. Microscopic mechanism for steady-state inhomogeneous flow in metallic glasses. *Acta Metall.* **25,** 407–415 (1977).
18. Schroers, J. & Johnson, W. L. Ductile bulk metallic glass. *Phys. Rev. Lett.* **93,** 255506 (2004).
19. Schroers, J., Lohwongwatana, B., Johnson, W. L. & Peker, A. Gold based bulk metallic glass. *Appl. Phys. Lett.* **87,** 061912 (2005).
20. Waniuk, T., Schroers, J. & Johnson, W. L. Timescales of crystallization and viscous flow of the bulk glass-forming Zr-Ti-Ni-Cu-Be alloys. *Phys. Rev. B* **67,** 184003 (2003).
21. Inoue, A., Shen, B. L., Koshiba, H., Kato, H. & Yavari, A. R. Cobalt-based bulk glassy alloy with ultrahigh strength and soft magnetic properties. *Nature Mater.* **2,** 661–663 (2003).
22. Chen, H. S. & Turnbull, D. Evidence of a glass-liquid transition in a gold-germanium-silicon alloy. *J. Chem. Phys.* **48,** 2560–2571 (1968).

23. Patterson, J. P. & Jones, D. R. H. Molding of a metallic glass. *Mater. Res. Bull.* **13,** 583–585 (1978).
24. Saotome, Y. & Inoue, A. in *7th IEEE Workshop on Micro Electro Mechanical Systems* 343–348 (IEEE, 1994).
25. Schroers, J. The superplastic forming of bulk metallic glasses. *J. Miner. Met. Mater.* **57,** 35–39 (2005).
26. Saotome, Y., Okaniwa, S., Kimura, H. & Inoue, A. Superplastic nanoforging of Pt-based metallic glass with dies of Zr-BMG and glassy carbon fabricated by focussed ion beam. *Mater. Sci. Forum* **539,** 2088–2093 (2007).
27. Schroers, J., Pham, Q., Peker, A., Paton, N. & Curtis, R. V. Blow molding of bulk metallic glass. *Scr. Mater.* **57,** 341–344 (2007).
28. Kumar, G. & Schroers, J. Write and erase mechanisms for bulk metallic glass. *Appl. Phys. Lett.* **92,** 031901 (2008).
29. Hua, F. *et al.* Polymer imprint lithography with molecular-scale resolution. *Nano Lett.* **4,** 2467–2471 (2004).
30. Dalla Torre, F., Van Swygenhoven, H., Schaublin, R., Spatig, P. & Victoria, M. Mechanical behaviour of nanocrystalline electrodeposited Ni above room temperature. *Scripta Mater.* **53,** 23–27 (2005).

Acknowledgements We thank T. A. Waniuk and C. S. O'Hern for reading of the manuscript and E. R. Dufresne for discussion. This work was supported by the US National Science Foundation under the Material Processing and Manufacturing programme.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.S. (jan.schroers@yale.edu).

METHODS

Custom heating plates were installed on a load cell of an Instron mechanical testing machine to allow a precise control of temperature and applied pressure during moulding experiments. The moulds (Si, Ni or Al_2O_3) were heated to 270 °C (for Pt-BMG) on the bottom heating plate of the modified Instron machine set-up. A piece of Pt-BMG was placed on the heated mould and after allowing 30 s for the temperature to equilibrate, the applied load was increased to attain a preset pressure value. The applied pressure was kept constant for varying time intervals depending on the mould type and features. For the moulding of high-aspect-ratio nanostructure, the Pt-BMG was formed on porous Al_2O_3 from 230 °C to 290 °C under a constant pressure of 130 MPa. The samples were quenched in water after removing from the load cell. Low-aspect-ratio structures separated following water quenching, owing to different thermal expansion coefficients of the mould and metallic glasses. High-aspect-ratio structures were separated by dissolving the mould in a KOH (40 wt%) solution heated to 80 °C.

The isothermal moulding resulted in uniform Pt-BMG structures that were subsequently used as moulds to imprint on Au-BMG and PMMA. The Au-BMG or PMMA was thermoplastically formed on patterned Pt-BMG at 160 °C. The Pt-BMG exhibits sufficient strength at 160 °C and there was no deterioration in the Pt-BMG following its use as a mould. In a separate experiment, the Pt-BMG was thermoplastically formed on porous Al_2O_3 at 280 °C for 2 min to crystallize the BMG. The crystallized Pt-BMG then served as a mould to imprint on the Pt-BMG at 270 °C. The partial crystallization was done in order to avoid the embrittlement of the BMG mould upon complete crystallization. For hot-cutting process, the BMGs formed on moulds were reheated in the supercooled liquid region and the BMG reservoir was scraped off the mould as described previously in detail[31]. The patterned BMGs were characterized by DSC (differential scanning calorimetry) and SEM.

31. Schroers, J., Pham, Q. & Desai, A. Thermoplastic forming of bulk metallic glass-A technology for MEMS and microstructure fabrication. *J. Microelectromech. Syst.* **16**, 240–247 (2007).

LETTERS

Magma-compensated crustal thinning in continental rift zones

H. Thybo[1] & C. A. Nielsen[1]

Continental rift zones are long, narrow tectonic depressions in the Earth's surface where the entire lithosphere has been modified in extension[1]. Rifting can eventually lead to rupture of the continental lithosphere and creation of new oceanic lithosphere or, alternatively, lead to formation of wide sedimentary basins around failed rift zones. Conventional models of rift zones include three characteristic features: surface manifestation as an elongated topographic trough, Moho shallowing due to crustal thinning, and reduced seismic velocity in the uppermost mantle due to decompression melting or heating from the Earth's interior[2–4]. Here we demonstrate that only the surface manifestation is observed at the Baikal rift zone, whereas the crustal and mantle characteristics can be ruled out by a new seismic profile across southern Lake Baikal in Siberia. Instead we observe a localized zone in the lower crust which has exceptionally high seismic velocity and is highly reflective. We suggest that the expected Moho uplift was compensated by magmatic intrusion into the lower crust, producing the observed high-velocity zone. This finding demonstrates a previously unknown role for magmatism in rifting processes with significant implications for estimation of stretching factors and modelling of sedimentary basins around failed rift structures.

The active Baikal rift zone (BRZ, Fig. 1) is located in the centre of the largest continental area on Earth, far from any active plate boundary. Its surface expression is a curved series of 40–80-km-wide graben structures (fault-bounded localized depressions), more than 2,000 km long, at the edge of the Siberian craton. The 660-km-long Lake Baikal in the central part of the rift is the deepest lake on Earth, and contains around 20% of the world's available freshwater resources. The surrounding elevated rift-shoulders reach altitudes of 1,500–2,500 m around the central rift (Lake Baikal is 455 m above sea level), which is similar to other active continental rifts. The BRZ developed during the last ~35 Myr in the Palaeozoic Sayan-Baikal accretionary belt at the late Ordovician-Silurian suture between the Siberian craton and the Amurian plate[5–7].

The development of the BRZ has produced only small volumes of volcanic products, most of which are observed at distances of >100 km from the graben structures[8] (Fig. 1). The total volume of volcanic rocks around the BRZ probably does not exceed 6,000 km^3, whereas more than 140,000 km^3 of volcanic rocks are estimated around the Kenya rift zone[9]. Weakly alkaline, basaltic volcanism at the BRZ began at around 40–35 Myr ago in the Tunka depression (Fig. 1) and was followed by volcanic pulses, primarily in the surrounding highlands[8].

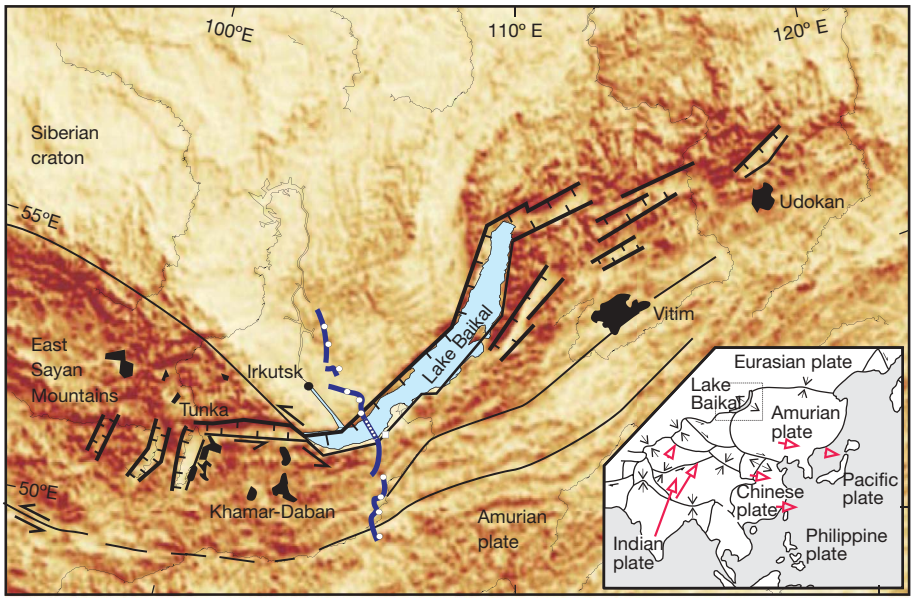


Figure 1 | Shaded relief topographic map and tectonic setting. Main figure, topographic map showing location of seismic profile (blue line) across the Baikal rift zone (BRZ). Seismic sources: large circles, explosions; small circles, airgun line; square, Babushkin supervibrator. Full lines, major tectonic faults; lines with tick marks, normal faults at main grabens[31]; black zones, volcanic provinces[15]. Inset, tectonic setting with the BRZ at the southeastern edge of the Siberian craton (Eurasian plate), modified from ref. 5. Rifting may be caused by eastward escape of terranes due to India–Eurasia collision, extension from Pacific subduction[13], and a possible mantle thermal anomaly[6,10]. Red arrows, GPS measurements of plate motion[32]; black arrows, relative displacement at plate boundaries.

[1]Department of Geography and Geology, University of Copenhagen, Oester Voldgade 10, DK-1350 Copenhagen K, Denmark.

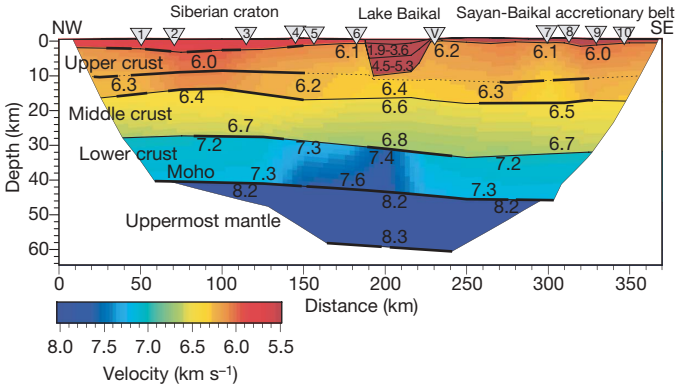


Figure 2 | Model of seismic compressional velocity along the seismic profile across the BRZ. Discontinuities are shown by lines, which are bold at observed reflection points (Methods); seismic velocity is colour-coded (scale at right). The Baikal rift depression is filled with low-velocity sediments ($<5.3\,km\,s^{-1}$). The crustal thickness is almost uniform and smoothly varying without Moho uplift. The mantle velocity is normal for the Siberian craton, and no low-velocity anomaly is observed to depths of 60 km (Fig. 3). A localized zone in the lower crust exhibits very high seismic velocity ($v > 7.4 \pm 0.2\,km\,s^{-1}$). This zone is strongly reflective (Fig. 3). Triangles with numbers mark shot point locations.

Extension rate in the BRZ has varied over time. Sedimentation rate corresponded to subsidence rate during the early phases of rift formation. It increased during regional uplift around 10–7 Myr ago[7], and Lake Baikal was probably created by fast subsidence since ~2.5 Myr ago. This sequence has been interpreted in terms of active rifting caused by a mantle plume[6,10], although the rifting was not preceded by significant volcanic activity[8]. Recent interpretation of teleseismic recordings indicates the presence of a narrow thermal anomaly in the mantle, which cannot be shallower than 70 km deep[11,12]. Alternatively, the timing and location of the rifting at Baikal can be explained by passive rifting processes[4] caused by the eastward escape of the Amurian plate (Fig. 1) due to the Himalayan continental collision and plate drag from the subduction of the Pacific plate underneath Asia[13–15].

We present a new seismic velocity model interpreted from a 360-km-long, crustal refraction seismic profile across southern Lake Baikal (Fig. 2), acquired by the Baikal Explosion Seismic Transects (BEST) project (Methods). The model is well constrained by reversed ray coverage along the whole profile. It provides a good fit to 3,290 out of 3,442 observed travel times with a root-mean-square residual of 103 ms, which corresponds to the uncertainty of the travel time picks. The ~10-km-deep and 40-km-wide graben structure underneath Lake Baikal is filled with sediments with seismic velocities (v) of $1.9–5.3\,km\,s^{-1}$. The ~10-km-thick sedimentary sequence may be rift-related, although the 4-km-thick lower interval ($v > 4.5\,km\,s^{-1}$) may also consist of pre-rift sedimentary rocks.

The gentle southward crustal thickening from 41 km in the Siberian craton to 46 km (±1 km) in the Sayan-Baikal accretionary belt implies no Moho uplift, contrary to indications from some low-resolution seismological studies[6,11,16] but in agreement with Deep

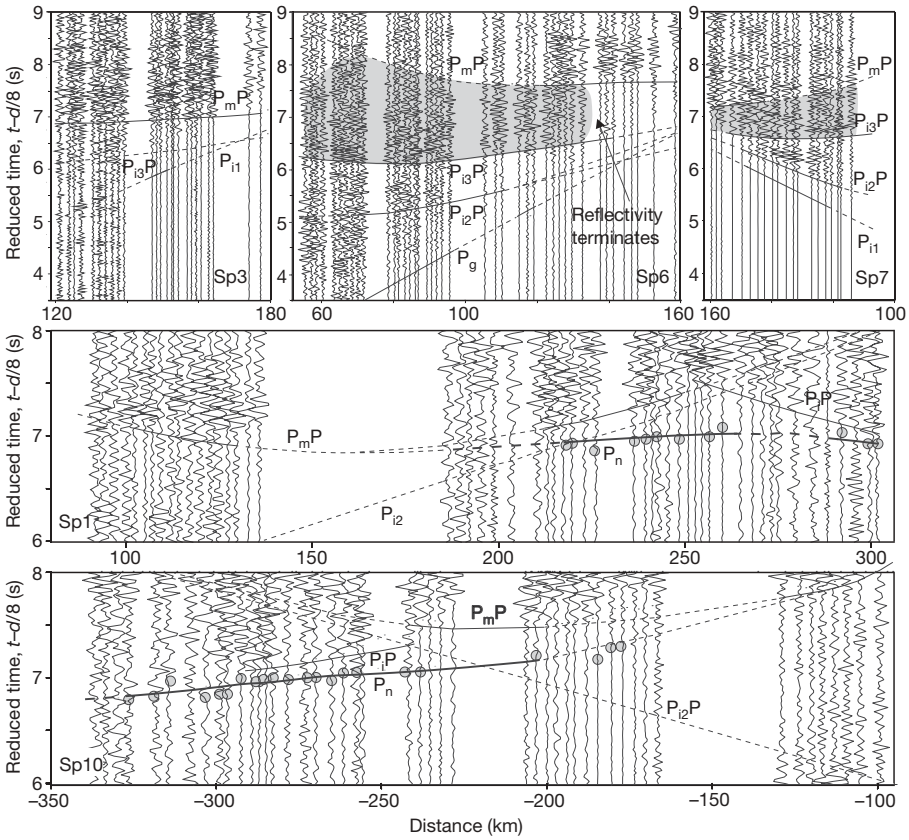


Figure 3 | Seismic sections across southern Lake Baikal with 8.0 km s^{-1} reduction velocity, with superimposed calculated travel times for the model in Fig. 2. The time scale has been reduced by subtracting distance divided by reduction velocity (v_r) from actual time. Full lines are shown where seismic phases are identified, otherwise dashed lines are used. Top left, Sp3 shows reflection from Moho below the rift zone. Lower-crustal reflectivity is modest owing to flat propagation angle (confirmed by full waveform synthetic seismograms). Top middle and right, Sp6 and Sp7 illustrate pronounced lower-crustal reflectivity (grey shading), restricted to a 50–80-km-wide zone at the BRZ. Terminations (as Sp6 reflectivity at ~130 km offset) constrain location of the reflective bodies to the high-velocity, reflective lower-crustal zone. Middle and bottom, Sp1 and Sp10 are reversed sections, demonstrating that seismic velocities are >8.0 km s^{-1} in the uppermost mantle along the whole profile. Circles show picked travel times. P_g and P_{i1}, crustal refractions; $P_{i2}P$ and $P_{i3}P$, intra-crustal reflections; P_mP, Moho reflection; P_n, refraction from the uppermost mantle; P_lP, reflection from ~60 km depth in the upper mantle (see also Supplementary Information).

Seismic Sounding results[17]. The crust–mantle transition, the Moho, is a reflector, where the velocity increases to $8.2 \pm 0.2\,km\,s^{-1}$ along the whole profile. No sign is observed of low seismic velocity in the uppermost mantle (Fig. 3). These remarkable observations contradict the existing models of rift formation, which predict Moho uplift beneath the rift for the pure shear model[2,6] or outside the rift for the simple shear model[3,14].

The lower crust includes a 50–80-km-wide, high-velocity anomaly 7.4–7.6 (± 0.2) $km\,s^{-1}$, beneath and slightly displaced to the northwest of the rift. Significant reverberant reflectivity originates from this zone (Fig. 3). We have modelled this reflectivity with a layered sequence of strong reflectors (layer thicknesses of 400 ± 150 m and large velocity contrasts of around $0.4\,km\,s^{-1}$ (Fig. 4a, Methods), which occupy around 50% of the 13-km-thick lower crust. The very large velocities and velocity contrasts indicate fractionation during cooling of the intrusions[18]. The estimated volume of intrusions ($0.5 \times 13 \times 50$–80 km) is equal to the volume of the sedimentary graben structure (10×40 km). As such, the presence of the intrusions explains the flat Moho because their volume compensates for the expected crustal thinning around the BRZ. The mass/area of vertical columns ($1.49 \times 10^8\,g\,m^{-2}$ to 50 km depth) is constant to within ~1% along the profile, as calculated for densities derived from the seismic velocity model. Hence, isostatic equilibrium of the crust is maintained as the low density of the sediments is compensated by the high density of the intrusions.

Seismic and gravity results[19] from the Palaeozoic Dniepr-Donetsk rift zone in Ukraine also show a flat Moho and reflectivity indicative of mafic intrusions in the lower crust with thicknesses similar to the BRZ. The Kenya rift also shows less Moho uplift than expected from the observed amount of stretching, which has been explained by the presence of intrusions with thicknesses of 200–500 m (ref. 18). The average velocities in the lower crust and uppermost mantle are lower below the Kenya rift than the BRZ, which may be due to high temperatures[20]. Reverberative seismic reflectivity below the rift zones may be caused by intermixed layering of primarily olivine- and pyroxene-rich rocks with layers of plagioclase-rich components. Numerical simulation shows that mafic intrusion into a cold crust will take the form of sills[21], in accordance with our observations.

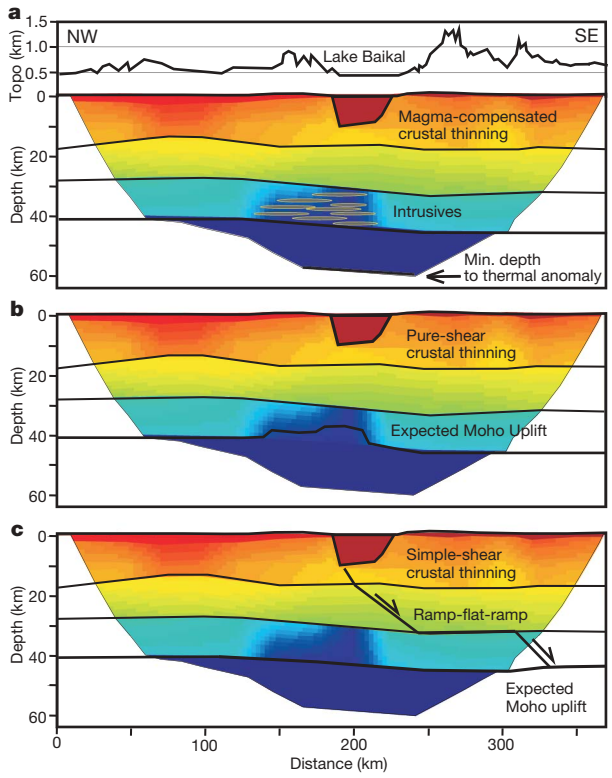


Figure 4 | Three models for formation of the BRZ superimposed on the seismic velocity model. **a**, Our preferred model in terms of magma-compensated crustal thinning: the high-velocity, reflective zone in the lower crust includes 50% of intrusive material, which explains the flat Moho. The deepest reflector indicates the minimum depth to the asthenosphere (or significant positive thermal anomaly).Topography is shown at top. **b**, Pure shear[2,4] predicts uplift of Moho below the BRZ (drawn as expected uplift without intrusions based on actual lower-crustal velocity), which is not observed in the seismic model. **c**, Simple shear[3] predicts uplift of Moho (here drawn according to ref. 14) outside the BRZ, which is not observed to a distance of 200 km. Moreover, this model does not explain the reflective zone in the lower crust with abnormally high seismic velocity.

Contrary to other active continental rift zones[1,22–24], there is no change in seismic velocity in the uppermost mantle beneath the BRZ. Hence, no thermal anomaly from the uppermost mantle has reached the base of the crust at the BRZ. Our results show that any low-velocity anomaly in the mantle must be located deeper than a distinct seismic reflector at 60 ± 5 km depth (Fig. 2), thus supporting a recent teleseismic determination of a >70-km-deep asthenosphere[11].

Our new results have implications for estimation of stretching factors for rift zones. The β-factor (the ratio between initial and final thickness of the crust) based on the total crustal thickness is 1.3, whereas it is 1.7 if the lower-crustal intrusions are taken into account (Fig. 4b). The β-factor and the local heating of the crust by intrusions are important parameters for modelling basin evolution and possible hydrocarbon formation around failed rift zones. A β-factor of 1.7 will lead to decompression melting of the mantle for a 100-km-thick lithosphere[25]. The present lithospheric thickness of 60–70 km (ref. 11) implies a similar stretching factor, which together with a slightly elevated geotherm[20] explains the melting. Petrologic studies indicate a slightly elevated mantle temperature at the depth of melting[15] and BRZ magmatism is similar to regional magmatism[26]. Hence a mantle thermal anomaly may be present but has not yet reached the surface[7].

The seismic velocities in the lower-crustal anomaly are higher than often observed at underplated continental margins. Some of the highest velocities observed at the North Atlantic margins[27,28] are in the interval 7.2–7.7 $km\,s^{-1}$, which corresponds to 7.4–7.9 $km\,s^{-1}$ at conditions in the BRZ lower crust (40 km depth and 400 °C temperature for[29] $dv/dp = 0.0002\,km\,s^{-1}\,MPa^{-1}$ and $dv/dT = -0.0004\,km\,s^{-1}\,°C^{-1}$, where p and v are respectively pressure and temperature). Theoretical velocities of oceanic crust[28,29] are 7.2–7.6 $km\,s^{-1}$ at BRZ conditions, slightly less than our observation. However, these predicted velocities may be underestimated for the BRZ because (1) the magma originates from a depleted continental mantle[15], (2) the magma may have fractionated during solidification[18], and (3) it solidified in a cold crust[20] at conditions favourable for metamorphism of gabbro into eclogitic rocks with a seismic velocity of 7.8–8.0 $km\,s^{-1}$ at a depth of 40 km, depending on the chemical composition[30]. Structural anisotropy due to layering of the intrusions will bias our observations towards the highest velocity instead of the average velocity, because the velocity observation is primarily based on horizontally travelling waves. Although very high, the observed velocity is within predicted values from laboratory measurements on mantle-derived magmatic rocks.

High-velocity lower crust has been earlier observed in a seismic profile along the strike of the BRZ and interpreted as normal cratonic lower crust with indication for simple shear deformation, where the top of the lower crust of the Siberian craton transforms the extension eastward to beneath the Mongolian plateau (Fig. 4c)[14]. In contrast, our new results with anomalously high velocity in a lower-crustal zone almost directly below the rift indicate instead mainly pure shear deformation (Fig. 4a). The intrusions may have formed during the slow rift phase (10–2.5 Myr ago).

The surface manifestation of the BRZ is typical for rift zones, but the seismic structure of the crust and uppermost mantle rules out the two other characteristic features of continental rift zones. Neither Moho uplift nor any low-velocity anomaly in the uppermost mantle are inferred from the new controlled source seismic profile across

southern Lake Baikal. These results indicate that no mantle thermal anomaly has reached close to the base of the crust, and that the rifting processes are not driven by a mantle plume. Significant volume of magmatic intrusions, balancing the expected crustal thinning, explains the apparent lack of crustal thinning beneath the BRZ.

METHODS SUMMARY

The seismic model is interpreted from data acquired by the Baikal Explosion Seismic Transects (BEST) project, which is a collaboration between the University of Copenhagen, the Siberian Branch of the Russian Academy of Sciences, and the Polish Academy of Sciences. BEST applied state-of-the-art refraction methods along the 360-km-long profile across southern Lake Baikal. Seismic sources included 10 chemical explosions in 50-m-deep boreholes (500–2,500 kg of TNT), 234 air gun shots (total volume 60 l) in Lake Baikal, and 6 vibration sweeps (5.8—9.7 Hz over 45 min) by the stationary Babushkin super-vibrator (100 ton weight). Seismic data were recorded by 175 Texan seismographs with 4.5 Hz vertical geophones. High quality signals are correlated along the whole profile with ± 100 ms uncertainty (± 300 ms for the supervibrator).

Data were interpreted by ray tracing in two-dimensional models of seismic velocity versus depth including discontinuities. Trial-and-error modelling (top-down approach) is based on observed travel times of refracted and reflected signals from the crust and uppermost mantle, followed by automatic inversion for selected parameters. Initial forward modelling is necessary because of the inherent nonlinearity of wave propagation. The model provides a good fit to 3,290 out of 3,442 observed travel times with root-mean-square residual travel times of 103 ms, which corresponds to the uncertainty of the picks. Ray-tracing synthetic seismograms improved the constraints on velocity gradients.

Strong, reverberating seismic reflectivity from a lower-crustal zone has been modelled by full waveform calculation of synthetic seismograms. The reflectivity method for one-dimensional models constrains average layer thicknesses to 400 ± 150 m and 0.4 km s^{-1} velocity contrasts. Calculations with the Tesseral program for two-dimensional models confirm reflectivity results and constrain the lateral extent of the reflective zone to coincide with the high-velocity zone in the lower crust below the BRZ (see also Supplementary Information).

Received 14 May; accepted 9 December 2008.

1. Olsen, K. H. (ed.) *Continental Rifts: Evolution, Structure, Tectonics* (Elsevier, 1995).
2. McKenzie, D. Some remarks on the development of sedimentary basins. *Earth Planet. Sci. Lett.* **40,** 25–32 (1978).
3. Wernicke, B. P. Uniform-sense normal simple shear of the continental lithosphere. *Can. J. Earth Sci.* **22,** 108–125 (1985).
4. Ruppel, C. Extensional processes in continental lithosphere. *J. Geophys. Res. B* **100,** 24,187–24,215 (1995).
5. Moore, T. C., Klitgord, K. D., Golmshtok, A. J. & Weber, E. Sedimentation and subsidence patterns in the central and north basins of Lake Baikal from seismic stratigraphy. *Geol. Soc. Am. Bull.* **109,** 746–766 (1997).
6. Zorin, Y. A. *et al.* The Baikal rift zone: The effect of mantle plumes on older structure. *Tectonophysics* **371,** 153–173 (2003).
7. Petit, C. & Deverchere, J. Structure and evolution of the Baikal rift: A synthesis. *Geochem. Geophys. Geosyst.* **7,** doi:10.1029/2006GC001265 (2006).
8. Kiselev, A. I. Volcanism of the Baikal rift zone. *Tectonophysics* **143,** 235–244 (1987).
9. Achauer, U. & Masson, F. Seismic tomography of continental rifts revisited: From relative to absolute heterogeneities. *Tectonophysics* **358,** 17–37 (2002).
10. Windley, B. F. & Allen, M. B. Mongolian plateau: Evidence for a late Cenozoic mantle plume under central Asia. *Geology* **21,** 295–298 (1993).
11. Petit, C., Koulakov, I. & Deverchere, J. Velocity structure around the Baikal rift zone from teleseismic and local earthquake traveltimes and geodynamic implications. *Tectonophysics* **296,** 125–144 (1998).
12. Zhao, D. P., Lei, J. S., Inoue, T., Yamada, A. & Gao, S. S. Deep structure and origin of the Baikal rift zone. *Earth Planet. Sci. Lett.* **243,** 681–691 (2006).
13. Tapponier, P. & Molnar, P. Active faulting and Cenezoic tectonics of the Tien Shan, Mongolia, and Baykal regions. *J. Geophys. Res.* **84,** 3425–3459 (1979).
14. ten Brink, U. S. & Taylor, M. H. Crustal structure of central Lake Baikal: Insights into intracontinental rifting. *J. Geophys. Res.* **107** (B7), doi:10.1029/2001JB000300 (2002).
15. Ionov, D. Mantle structure and rifting processes in the Baikal-Mongolia region: Geophysical data and evidence from xenoliths in volcanic rocks. *Tectonophysics* **351,** 41–60 (2002).
16. Gao, S. S., Liu, K. H. & Chen, C. Significant crustal thinning beneath the Baikal rift zone: New constraints from receiver function analysis. *Geophys. Res. Lett.* **31,** L20610 (2004).
17. Suvorov, V. D. *et al.* Structure of the crust in the Baikal rift zone and adjacent areas from Deep Seismic Sounding data. *Tectonophysics* **351,** 61–74 (2002).
18. Thybo, H., Maguire, P. K. H., Birt, C. & Perchuc, E. Seismic reflectivity and magmatic underplating beneath the Kenya Rift. *Geophys. Res. Lett.* **27,** 2745–2748 (2000).
19. Lyngsie, S. B., Thybo, H. & Lang, R. Rifting and lower crustal reflectivity: A case study of the intracratonic Dniepr-Donets rift zone, Ukraine. *J. Geophys. Res.* **112,** B12402, doi:10.1029/2006JB004795 (2007).
20. Artemieva, I. M. & Mooney, W. D. Thermal thickness and evolution of Precambrian lithosphere: A global study. *J. Geophys. Res.* **106** (B8), 16387–16414 (2001).
21. Gerya, T. V. & Burg, J. P. Intrusion of ultramafic magmatic bodies into the continental crust: Numerical simulation. *Phys. Earth Planet. Inter.* **160,** 124–142 (2007).
22. Prodehl, C. *et al.* Large-scale variation in lithospheric structure along and across the Kenya Rift. *Nature* **354,** 223–227 (1991).
23. Green, W. V., Achauer, U. & Meyer, R. P. A three-dimensional seismic image of the crust and upper mantle beneath the Kenya Rift. *Nature* **354,** 199–203 (1991).
24. Wilson, D. *et al.* Lithospheric structure of the Rio Grande rift. *Nature* **433,** 851–855 (2005).
25. White, R. & McKenzie, D. Magmatism at rift zones — the generation of volcanic continental margins and flood basalts. *J. Geophys. Res.* **94** (B6), 7685–7729 (1989).
26. Barry, T. L. *et al.* Petrogenesis of Cenozoic basalts from Mongolia: Evidence for the role of asthenospheric versus metasomatized lithospheric mantle sources. *J. Petrol.* **44,** 55–91 (2003).
27. Raum, T. *et al.* Crustal structure and evolution of the southern Voring basin and Voring transform margin, NE Atlantic. *Tectonophysics* **415,** 167–202 (2006).
28. White, R. S., Smith, L. K., Roberts, A. W., Christie, P. A. F. & Kusznir, N. J. Lower-crustal intrusion on the North Atlantic continental margin. *Nature* **452,** 460–464 (2008).
29. Korenaga, J., Kelemen, P. B. & Holbrook, W. S. Methods for resolving the origin of large igneous provinces from crustal seismology. *J. Geophys. Res.* **107** (B9), doi:10.1029/2001JB001030 (2002).
30. Furlong, K. P. & Fountain, D. M. Continental crustal underplating — thermal considerations and seismic-petrologic consequences. *J. Geophys. Res.* **91** (B8), 8285–8294 (1986).
31. Delvaux, D. *et al.* Paleostress reconstructions and geodynamics of the Baikal region, Central Asia. Part 2. Cenozoic rifting. *Tectonophysics* **282,** 1–38 (1997).
32. Jin, S. G., Park, P. H. & Zhu, W. Y. Micro-plate tectonics and kinematics in Northeast Asia inferred from a dense set of GPS observations. *Earth Planet. Sci. Lett.* **257,** 486–496 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This study received support from the Carlsberg Foundation and the Danish Natural Science Research Council. The field work at Lake Baikal further received support from the Russian Academy of Sciences, Siberian Branch, and the Polish Academy of Sciences. The seismic instruments were provided by the University of Copenhagen and the Technical University of Vienna. We acknowledge discussions with R. S. White on melting processes, and comments received from I. Artemieva, I. Reid and W. Stratford on earlier versions of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to H.T. (thybo@geo.ku.dk).

LETTERS

A burst of segmental duplications in the genome of the African great ape ancestor

Tomas Marques-Bonet[1,2], Jeffrey M. Kidd[1], Mario Ventura[3], Tina A. Graves[4], Ze Cheng[1], LaDeana W. Hillier[4], Zhaoshi Jiang[1], Carl Baker[1], Ray Malfavon-Borja[1], Lucinda A. Fulton[4], Can Alkan[1], Gozde Aksay[1], Santhosh Girirajan[1], Priscillia Siswara[1], Lin Chen[1], Maria Francesca Cardone[3], Arcadi Navarro[2,5], Elaine R. Mardis[4], Richard K. Wilson[4] & Evan E. Eichler[1]

It is generally accepted that the extent of phenotypic change between human and great apes is dissonant with the rate of molecular change[1]. Between these two groups, proteins are virtually identical[1,2], cytogenetically there are few rearrangements that distinguish ape–human chromosomes[3], and rates of single-base-pair change[4–7] and retrotransposon activity[8–10] have slowed particularly within hominid lineages when compared to rodents or monkeys. Studies of gene family evolution indicate that gene loss and gain are enriched within the primate lineage[11,12]. Here, we perform a systematic analysis of duplication content of four primate genomes (macaque, orang-utan, chimpanzee and human) in an effort to understand the pattern and rates of genomic duplication during hominid evolution. We find that the ancestral branch leading to human and African great apes shows the most significant increase in duplication activity both in terms of base pairs and in terms of events. This duplication acceleration within the ancestral species is significant when compared to lineage-specific rate estimates even after accounting for copy-number polymorphism and homoplasy. We discover striking examples of recurrent and independent gene-containing duplications within the gorilla and chimpanzee that are absent in the human lineage. Our results suggest that the evolutionary properties of copy-number mutation differ significantly from other forms of genetic mutation and, in contrast to the hominid slowdown of single-base-pair mutations, there has been a genomic burst of duplication activity at this period during human evolution.

We began by developing a segmental duplication map for each of the four primate genomes (macaque, orang-utan, chimpanzee and human; Supplementary Fig. 1). The approach is based on the alignment of whole-genome shotgun (WGS) sequence data against the human reference genome and predicts high-identity segmental duplications based on excess depth of coverage and sequence divergence[13] (Methods). Previous analyses have suggested excellent sensitivity and specificity for computational detection of duplications larger than 20 kilobases (kb) in length[13] (Table 1, Supplementary Table 1 and Supplementary Note Table 2). By this criterion, we characterized 73 megabases (Mb) corresponding to the duplications identified in at least one of the four primate species, correcting for copy number in each primate (Methods). Furthermore, we characterized each duplication as 'lineage specific' or 'shared', depending on whether it was seen in only one or multiple genomes. This comparative map (Supplementary Figs 3 and 4) is available as an interactive UCSC mirror browser (http://humanparalogy.gs.washington.edu), allowing researchers to interrogate the evolutionary history of any duplicated region of interest.

We validated our primate genomic duplication map using two different experimental approaches and, wherever possible, using DNA from the same individuals from which the computational predictions were generated. Using fluorescence *in situ* hybridization (FISH), we found that 86.5% of segmental duplications were concordant with computational predictions when categorized as either lineage specific (50 out of 58) or shared duplications (40 out of 46) (Supplementary Figs 1 and 2; see also Fig. 1 and Supplementary Tables 2–4). As a second approach, we designed a specialized oligonucleotide microarray (1 probe per 585 bp) targeted to primate segmental duplications

Table 1 | Classes of primate segmental duplication

Category	Segmental duplications	Segmental duplications >20 kb	Validation (%)	Copy-number-corrected duplicated base pairs			
				Hsa	Ptr	Ppy	Mmu
Hsa	51,458,805	15,236,422	89–92	17,847,869	–	–	–
Ptr	11,239,390	4,789,874	99	–	16,583,946	–	–
Ppy	30,553,228	6,417,679	98	–	–	23,327,737	–
Mmu	24,962,092	5,360,646	45	–	–	–	45,810,964
Mmu*	35,493,466	7,715,410	85	–	–	–	18,266,656
Hsa/Ptr	32,392,480	21,061,194	NA	21,524,417	26,304,286	–	–
Hsa/Ptr/Ppy	25,450,827	13,402,545	NA	11,259,061	14,012,351	11,541,148	–
Hsa/Ptr/Ppy/Mmu	14,094,156	7,156,616	NA	8,092,997	12,820,607	6,176,876	12,542,691
Total	190,150,978	73,424,976	–	58,724,344	69,721,190	41,045,761	30,809,347

Duplications were divided into eight categories based on the WSSD analysis of each primate genome (subsequent analyses were restricted to segmental duplications >20 kb in length). Lineage-specific and shared duplication content are indicated. Percentage validation indicates the fraction of species-specific duplications confirmed by cross-species array comparative genomic hybridization. Because the human genome was used, we corrected for copy number and examined sequence contigs not aligned to the human genome (see Methods). Segmental duplications assigned to the Y chromosome were not considered.

* Macaque segmental duplications detected in the macaque reference genome using WSSD and WGAC (<94% identity) approaches.

[1]Department of Genome Sciences, University of Washington and the Howard Hughes Medical Institute, Seattle, Washington 98195, USA. [2]Institut de Biologia Evolutiva (UPF-CSIC), 08003 Barcelona, Catalonia, Spain. [3]Sezione di Genetica-Dipartimento di Anatomia Patologica e Genetica, University of Bari, 70125 Bari, Italy. [4]Genome Sequencing Center, Washington University School of Medicine, St Louis, Missouri 63108, USA. [5]Institució Catalana de Recerca i Estudis Avançats (ICREA) and Instituto Nacional de Bioinformática (INB), Dr. Aiguader 88, 08003 Barcelona, Spain.

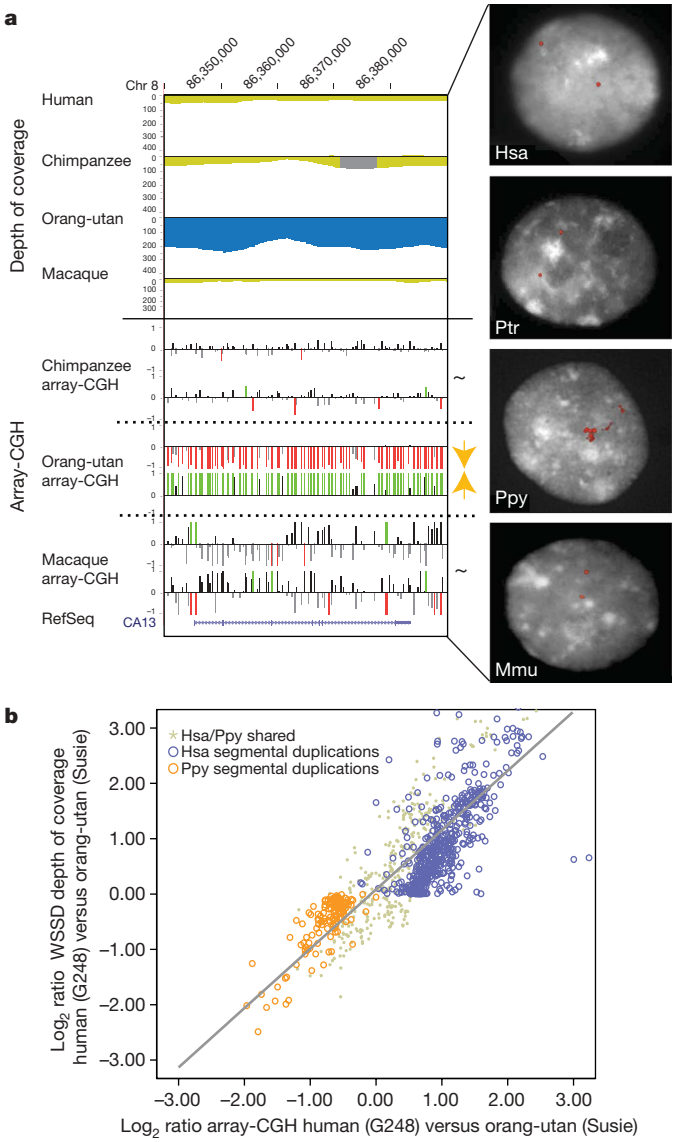


Figure 1 | Experimental validation of duplication map. **a**, A computationally predicted orang-utan-specific duplication (blue, excess depth of coverage of aligned WGS sequence) is confirmed by interspecific FISH and array-CGH (oligonucleotide relative $\log_2$ ratios are depicted as red/green histograms and correspond to an increase and decrease in signal intensity when test/reference are reverse labelled; see Supplementary Note for additional details). Hsa, *Homo sapiens* (human); Mmu, *Macaca mulatta* (macaque); Ptr, *Pan troglodytes* (chimpanzee); Ppy, *Pongo pygmaeus* (orang-utan). **b**, A comparison of duplication copy number as predicted by WSSD sequence analysis versus oligonucleotide array-CGH between nonhuman and human primates showed a good correlation ($r = 0.77$). Relative duplication copy number was computed as the $\log_2$ ratio of the number of aligned nonhuman primate reads against the human reference genome over the number of reads mapping to known single-copy BACs.

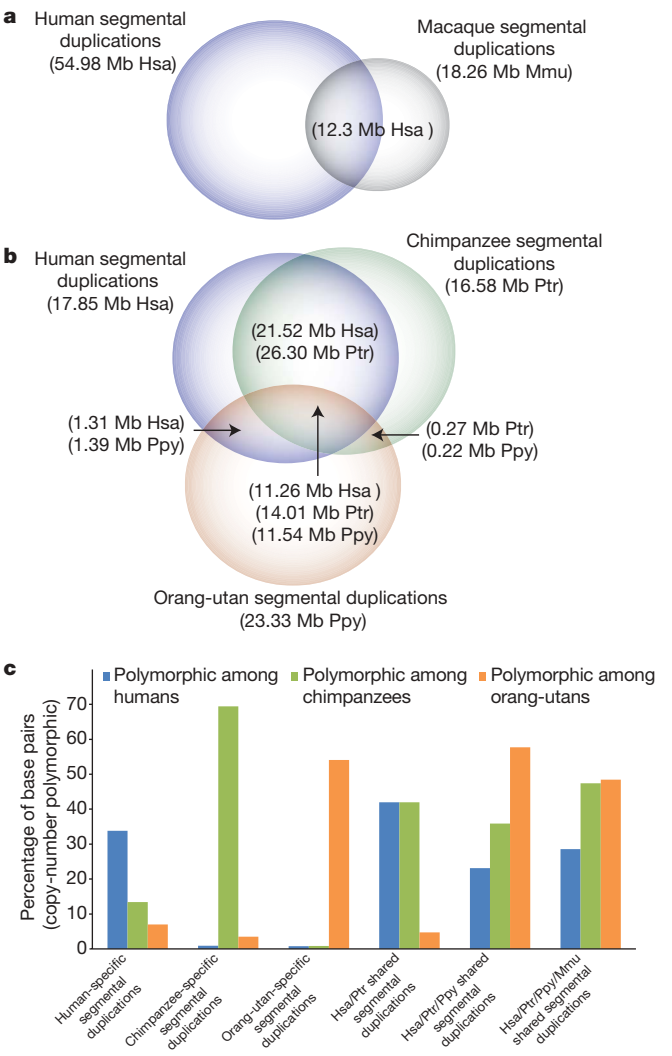


Figure 2 | Shared versus lineage-specific duplications and great-ape polymorphism. Segmental duplications (>20 kb) were classified as lineage specific or shared based on a four-way comparison of human, chimpanzee, orang-utan and macaque genomes. **a**, Human segmental duplications were compared to Old World monkey segmental duplications (based on a separate analysis of the macaque assembly[9]). **b**, As nonhuman great-ape duplications were detected based on alignment of WGS sequence against the human genome, we corrected for copy number based on the depth of coverage of WGS sequence (in brackets with the name of the species for which the correction was performed, see Table 1). **c**, Copy-number polymorphic regions were estimated from the results of array-CGH hybridizations between eight samples each of human, chimpanzee and orang-utan (using the same reference as the computational prediction). The proportion of duplicated bases that showed evidence of copy-number polymorphism (that is, gain or loss for ≥two individuals within each species) is shown for each class of segmental duplication (>20 kb).

(Table 1) and performed array comparative genomic hybridization (array-CGH) between species (Table 1, Fig. 1 and Supplementary Figs 2–4). Among the great-ape genomes, we confirmed 89–99% of the lineage-specific duplications by interspecific array-CGH (Table 1) with a very good correlation between computationally predicted and experimentally validated copy-number differences (Fig. 1b). Because only 45% of macaque-specific duplications could be confirmed by interspecific array-CGH, we performed an independent assessment of the macaque genome assembly and conservatively validated ~85% of macaque-specific duplications[9,14] (Z.J. and E.E.E., unpublished results).

The comparative duplication map reveals several important features of primate segmental duplications. As expected, most (80% or ~55 Mb) high-identity human segmental duplications arose after the divergence of the Old World monkey and hominoid lineages (Fig. 2a). Humans and chimpanzees show significantly more duplications than either macaque or orang-utan (Fig. 2b), with a large fraction being shared between chimpanzee and human. On the basis of our four-way primate genome analysis and leveraging array-CGH data from gorilla and bonobo (*Pan paniscus*), we classify only ~10 Mb of duplication content as human specific (210 duplication intervals with an average length of 53.1 kb). The genomic distribution of great-ape segmental duplications is highly nonrandom (Supplementary Fig. 5), with the presence of ancestral duplications being a strong predictor of 'new', lineage-specific events ($P < 0.001$, randomization test,

Supplementary Note Table 5a, b). For example, 45% of human–chimpanzee shared duplications map within 5 kb of segmental duplications shared among human–chimpanzee–orang-utan, whereas 31% of human–chimpanzee–orang-utan duplications map adjacent to human–chimpanzee–orang-utan–macaque duplications. These observations emphasize that unique sequences flanking more ancient duplications have a much higher probability of segmental duplication[13,15] and the duplication process itself is not random.

Within the human-specific set of duplications, we identified 39 partial and 17 complete human genes (Supplementary Table 7). As expected, we found that full-length hominid genes show greater evidence of positive selection when compared with similarly analysed unique genes (Supplementary Note). Our analysis indicates that several genes associated with human adaptation (amylase (*AMY1*), aquaporin 7 and *NBPF15*) are shared with chimpanzee but humans show a general increase in copy number. Gene models associated with signal transduction, neuronal activities (for example, neurotransmitter release, synaptic transmission) and muscle contraction are significantly enriched in human, chimpanzee and orang-utan lineage-specific duplications (Supplementary Table 7). Human and great-ape shared duplications or those shared with macaque are, in contrast, enriched for biological processes associated with amino acid metabolism ($P = 1.69 \times 10^{-2}$; great-ape shared segmental duplications) or oncogenesis ($P = 5.80 \times 10^{-13}$, 4.64×10^{-6}; ape segmental duplications shared with macaque). Although the number of such duplication events is few, these data suggest a shift in the types of genes that have been duplicated most recently during great-ape and human evolution.

There are two important caveats to the above analysis. First, we have analysed a single individual in each case and it is unclear to what extent that single genome represents the duplication pattern of the species. Second, duplicated sequences shared by two or more species might have potentially been subjected to recurrent mutations (homoplasy) leading to an overestimate of the proportion of ancestral duplications. Both copy-number polymorphism and recurrent duplication, in principle, will complicate classification of segmental duplications as 'ancestral' or 'lineage specific'. We therefore performed a number of additional analyses to address the impact of polymorphism and recurrent events on our assignments.

First, we investigated the extent of copy-number variation for both shared and lineage-specific duplications. Using array-CGH targeted to primate segmental duplications, we assessed the extent of copy-number variation in a set of unrelated DNA samples (Fig. 2c; see Methods). As expected[16,17], lineage-specific segmental duplications are highly copy-number variant, with humans showing 1.5- to two-fold less diversity in copy number when compared to chimpanzees and orang-utans (Fig. 2c; see also Supplementary Note Table 9). Notably, we found that shared segmental duplications are as copy-number variant as lineage-specific duplications and that humans show slightly greater copy-number variation for these (42% versus 34%) when compared with great apes.

It is, however, important to distinguish between duplication copy-number variation versus duplication status. A segmental duplication may show a high level of copy-number variation whereas its status as duplicated remains relatively constant among different individuals within a species. To address this, we performed a series of three-way array-CGH comparisons (Supplementary Note Fig. 7; see also Methods) where we investigated how duplication status (human-specific, chimpanzee-specific and orang-utan-specific segmental duplications) varied as function of copy-number polymorphism within a species. The results from these triangulations indicate that only 1–8% of the segmental duplications change duplication status even though 18–32% of the duplications are copy-number polymorphic between two

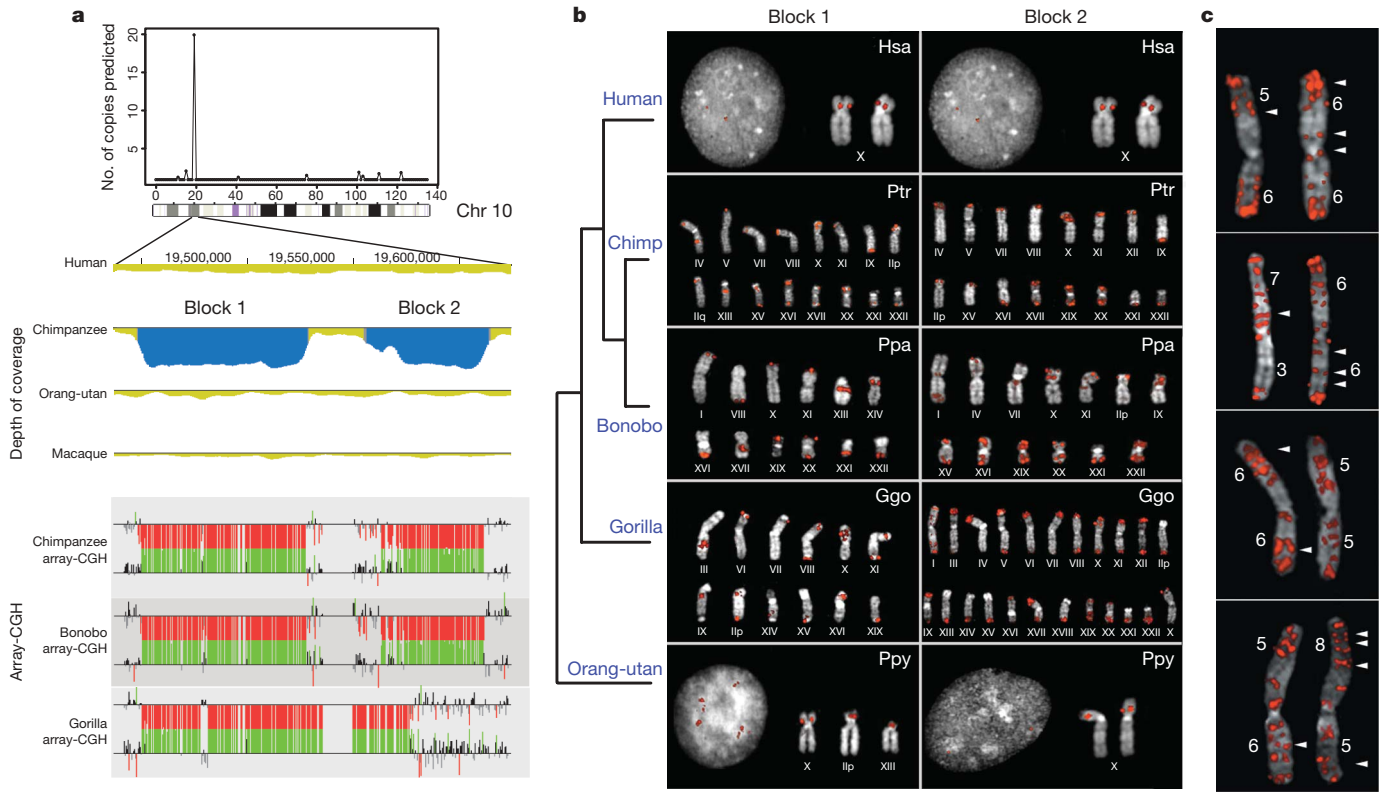


Figure 3 | Convergent gene duplication expansion in African great apes but not humans. **a**, Two regions on chromosome 10 have expanded in chimpanzee, gorilla and bonobo when compared to human based on computational and interspecific array-CGH experiments (see Fig. 1 legend). **b**, FISH confirms 23–50 copies in chimpanzee and bonobo (Ppa, *Pan paniscus*), and >100 copies in gorilla (Ggo, *Gorilla gorilla*) (Methods). End-sequence pair analysis using gorilla and chimpanzee WGS sequences reveals that all but the ancestral location are non-orthologous, indicating independent expansions in both lineages. **c**, Detailed analysis of eight homologues of gorilla chromosome 1 reveals interstitial locations of the block 2 duplication that show variation both in copy number and in terms of location.

individuals within a species (Supplementary Note Fig. 8). As a second independent test, we compared the duplication maps of two human genomes (J. C. Venter or HuRef and J. D. Watson genomes)[18,19] and found that 89% (595 out of 666) of the regions are shared duplications between HuRef and the J. D. Watson genome. Although we predict copy-number differences between these shared duplications, the boundaries of the duplication intervals remain remarkably consistent (Supplementary Fig. 7), suggesting again that duplication status is a relatively constant character state within a species.

To assess the potential impact of recurrent mutations leading to misclassification of ancestral events, we focused on shared duplications between human and chimpanzee that were not identified as duplicated in either orang-utan or macaque. We examined 103 sets of chimpanzee–human shared duplications that mapped to two or more distinct locations in the human genome (Supplementary Note) and determined what fraction of these mapped to two or more orthologous positions between chimpanzee and human. Using a paired end-sequence mapping approach[20,21] (Supplementary Note Fig. 9), we found that 85% (88 out of 103) of the chimpanzee–human shared duplications have two or more copies mapping to the same orthologous position in the two genomes. This implies that most of the shared duplications were already duplicated in the human–chimpanzee common ancestor (Supplementary Note Tables 6 and 7).

As part of our comparative analyses, we identified regions for which duplication patterns were inconsistent with the generally accepted human/great-ape phylogeny (Supplementary Fig. 4 and Supplementary Tables 5 and 6). For example, we identified 43 intervals that are duplicated in human and gorilla but not chimpanzee ($H^+C^-G^+$ duplications). Such a scenario may arise as a result of a deletion event in the chimpanzee lineage, incomplete lineage sorting or, less likely, recurrent duplication events in the human and gorilla lineages. Only the latter possibility would potentially lead to an overestimation of ancestral duplication events. We estimated the frequency of such events by mapping the location of the duplications in each species using paired end-sequence data[21] (see Supplementary Note). If the duplicated sequence mapped to the same location in gorilla and human, we classified it as a chimpanzee-specific deletion event or incomplete lineage sorting. If mapping to different locations in the two genomes, we categorized it as a recurrent event. As expected, most of the informative $H^+C^-G^+$ duplications (80% or 12 out of 15) were the result of chimpanzee-specific deletions.

We investigated the most extreme example of recurrent African ape duplications in more detail (Fig. 3). We identified a region (~150 kb in length) mapping to human chromosome 10 that had expanded in the chimpanzee genome but was largely single copy in human and orang-utan. It consists of two distinct duplication blocks (~86 and 66 kb in length). Both array-CGH and fluorescent *in situ* hybridization (FISH; Fig. 3a, b) confirm that the segments had been duplicated multiple times (~5–100 copies depending on the block and species) in the chimpanzee, bonobo and gorilla genomes but are single copy in all humans tested. Notably, the duplication boundaries (as delimited by array-CGH) differ between the gorilla and chimpanzee lineages. With the exception of the chromosome 10 locus, we found that the map locations between gorilla and chimpanzee are non-orthologous (Supplementary Note and Methods), indicating that this duplication expansion has occurred independently in both lineages.

On the basis of the large number of interstitial sites on gorilla chromosomes, we compared chromosome 1 from four unrelated gorillas for variation in copy number and location of this segmental duplication. Remarkably, we found that both copy number (10–14 copies per homologous chromosome) as well as map location for this segmental duplication vary among these eight gorilla homologues with as many as 50% of the map locations being unoccupied by a duplication in another homologue (Fig. 3c and Supplementary Fig. 13). We conclude that this ancestral region of chromosome 10 has served as a preferred donor of chimpanzee/great-ape duplications and that the chimpanzee and gorilla genomes have been restructured by independent bursts of duplication activity. Interestingly, we detected and confirmed by RT–PCR (reverse transcription PCR) at least one previously uncharacterized gene (14 exons, 141 kb of genomic sequence, 1,311 nucleotides of coding sequence and 437 amino acids) mapping to duplication block 1, which shows significant similarity to endosomal glycoprotein genes (Supplementary Note Figs 14–17). Thus, these duplications, in principle, may have led to African great-ape gene family expansions while remaining conspicuously a single copy in the human lineage. Although the mechanism by which such events have occurred is unclear, our data highlight the rapidity by which segmental duplications have restructured hominid genomes and emphasize their nonrandom nature both temporally and spatially.

Based on our genome-wide assessment of segmental duplications in each of four primate species and our estimate of 20% homoplasy

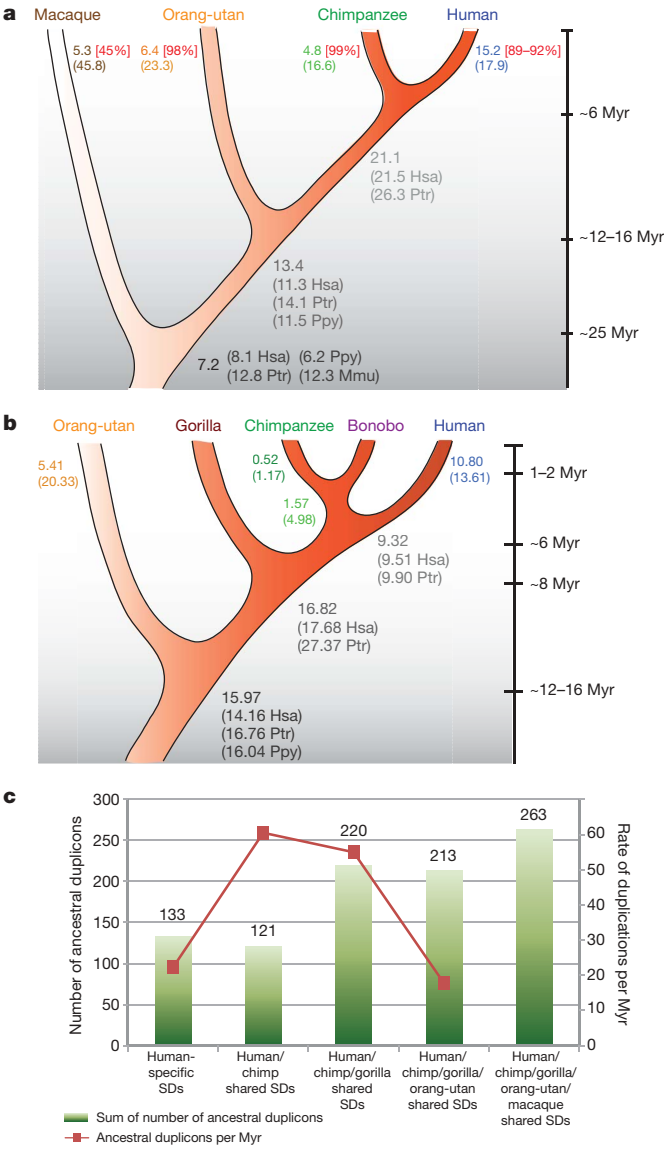


Figure 4 | Rates of segmental duplication. **a**, By base pair: we parsimoniously assigned the number of megabases to different branches, correcting for copy number in each species (shown in brackets). 89–99% of great-ape segmental duplications (SDs) were validated by array-CGH (square brackets). **b**, Further categorization of the segmental duplication data, based on array-CGH from bonobo and gorilla, shows the greatest accumulation in the ancestor of humans and great apes. **c**, By event: we assigned 950 evolutionarily distinct human segmental duplication events[22] to the human/great-ape phylogeny based on array-CGH results. The red line estimates the duplication rate (per million years (Myr)) and suggests an excess of large duplications in the common ancestor of human and chimpanzee but after the separation from gorilla.

(see above), we calculated rates of segmental duplication both in events[22] and base pairs along each lineage and ancestral node (Fig. 4, Supplementary Note Tables 13–16). We developed a maximum likelihood model to test if the rate of accumulation of segmental duplication has remained constant during the course of human/great-ape evolution. We compared the likelihood that the rate of segmental duplication has been uniform versus the likelihood of differential rates within specific lineages (Fig. 4). We find a significant increase (likelihood ratio test (LRT), $P < 1 \times 10^{-10}$) in both the number of events and base pairs in the human/African great-ape lineage when compared to macaque/Old World monkey lineage. Although terminal hominid lineages show an excess of duplications, the most significant burst of activity (4–10-fold, LRT $P = 1 \times 10^{-10}$) occurs in the common ancestor of human/chimpanzee and gorilla and after divergence of gorilla from the human–chimpanzee lineage (Supplementary Note Table 17). Our prediction is in strong agreement with the degree of sequence divergence among human intrachromosomal segmental duplications that shows a mode at 97–99% sequence identity. We note that this burst of duplication activity corresponds to a time when other mutational processes, such as point substitutions and retrotransposon activity, were slowing along the hominoid lineage. This apparent burst of activity may be the result of changes in the effective population size, generation time or imply a genomic destabilization at a period before and perhaps during hominid speciation. In light of the importance of segmental duplications in contributing to copy-number changes associated with neurocognitive disease[23–26] and disease susceptibility[27–29], we predict that this apparent acceleration has had a profound impact on the reproductive success, adaptability and evolution of ancestral hominid populations.

METHODS SUMMARY

We estimated the duplication content of human, chimpanzee, orang-utan and macaque by the whole-genome shotgun sequence detection (WSSD) method[13,30]. We mapped high-quality whole-genome shotgun (WGS) sequence reads for all species against the human reference assembly (NCBI build35) and identified regions of excess depth of coverage and divergence (see Supplementary Note). We also mapped macaque WGS reads to the macaque assembly (v 1.0). In this analysis, we considered segmental duplications >20 kb and >94% of identity (88% of identity for macaque reads against the human genome). We used read depth to estimate the number of copies for each duplication due to the excellent correlation ($r^2 = 0.953$)[13] between probes of known copy number and WGS depth of coverage.

We constructed an oligonucleotide microarray ($n = 385{,}000$) targeted to regions of primate segmental duplication (~180 Mb) and performed cross-species array-CGH (with human as a reference; GEO accession number GSE13884). With the exception of human, we used DNA derived from the same genome that was sequenced as part of primate genome sequencing projects. The same microarray was used to assess copy-number polymorphism in DNA samples from eight humans, eight chimpanzees and eight orang-utans (GSE13885). We also used FISH to validate further a subset of our duplications among the great apes.

We used end-sequence pair data from fosmid clones from a single human and a single chimpanzee as well as plasmid clones from a gorilla to map the location of segmental duplications within great-ape genomes (sequence data available from NIH trace repository). To estimate rates of segmental duplication along the hominoid phylogeny, we modelled the accumulation of segmental duplications in each branch as a pure birth process within a maximum likelihood framework. Nested models of segmental duplication were tested against each other by means of likelihood ratio tests (Supplementary Note).

Received 29 August; accepted 18 December 2008.

1. King, M. C. & Wilson, A. C. Evolution at two levels in humans and chimpanzees. *Science* **188,** 107–116 (1975).
2. Goodman, M. The role of immunochemical differences in the phyletic development of human behavior. *Hum. Biol.* **33,** 131–162 (1961).
3. Yunis, J. J. & Prakash, O. The origin of man: a chromosomal pictorial legacy. *Science* **215,** 1525–1530 (1982).
4. Wu, C. I. & Li, W. H. Evidence for higher rates of nucleotide substitution in rodents than in man. *Proc. Natl Acad. Sci. USA* **82,** 1741–1745 (1985).
5. Li, W. H. & Tanimura, M. The molecular clock runs more slowly in man than in apes and monkeys. *Nature* **326,** 93–96 (1987).
6. Elango, N., Thomas, J. W. & Yi, S. V. Variable molecular clocks in hominoids. *Proc. Natl Acad. Sci. USA* **103,** 1370–1375 (2006).
7. Steiper, M. E., Young, N. M. & Sukarna, T. Y. Genomic data support the hominoid slowdown and an Early Oligocene estimate for the hominoid-cercopithecoid divergence. *Proc. Natl Acad. Sci. USA* **101,** 17021–17026 (2004).
8. Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420,** 520–562 (2002).
9. Rhesus Macaque Genome Sequencing and Analysis Consortium. Evolutionary and biomedical insights from the rhesus macaque genome. *Science* **316,** 222–234 (2007).
10. The Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437,** 69–87 (2005).
11. Hahn, M. W., Demuth, J. P. & Han, S. G. Accelerated rate of gene gain and loss in primates. *Genetics* **177,** 1941–1949 (2007).
12. Dumas, L. *et al.* Gene copy number variation spanning 60 million years of human and primate evolution. *Genome Res.* **17,** 1266–1277 (2007).
13. Cheng, Z. *et al.* A genome-wide comparison of recent chimpanzee and human segmental duplications. *Nature* **437,** 88–93 (2005).
14. Jiang, Z., Hubley, R., Smit, A. & Eichler, E. E. DupMasker: A tool for annotating primate segmental duplications. *Genome Res.* **18,** 1362–1368 (2008).
15. Stankiewicz, P., Shaw, C. J., Withers, M., Inoue, K. & Lupski, J. R. Serial segmental duplications during primate evolution result in complex human genome architecture. *Genome Res.* **14,** 2209–2220 (2004).
16. Perry, G. H. *et al.* Hotspots for copy number variation in chimpanzees and humans. *Proc. Natl Acad. Sci. USA* **103,** 8006–8011 (2006).
17. Lee, A. S. *et al.* Analysis of copy number variation in the rhesus macaque genome identifies candidate loci for evolutionary and human disease studies. *Hum. Mol. Genet.* **17,** 1127–1136 (2008).
18. Levy, S. *et al.* The diploid genome sequence of an individual human. *PLoS Biol.* **5,** e254 (2007).
19. Wheeler, D. A. *et al.* The complete genome of an individual by massively parallel DNA sequencing. *Nature* **452,** 872–876 (2008).
20. Tuzun, E. *et al.* Fine-scale structural variation of the human genome. *Nature Genet.* **37,** 727–732 (2005).
21. Newman, T. L. *et al.* A genome-wide survey of structural variation between human and chimpanzee. *Genome Res.* **15,** 1344–1356 (2005).
22. Jiang, Z. *et al.* Ancestral reconstruction of segmental duplications reveals punctuated cores of human genome evolution. *Nature Genet.* **39,** 1361–1368 (2007).
23. Lee, J. A. & Lupski, J. R. Genomic rearrangements and gene copy-number alterations as a cause of nervous system disorders. *Neuron* **52,** 103–121 (2006).
24. Sharp, A. J. *et al.* Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nature Genet.* **38,** 1038–1042 (2006).
25. The International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* **455,** 237–241 (2008).
26. Sebat, J. *et al.* Strong association of *de novo* copy number mutations with autism. *Science* **316,** 445–449 (2007).
27. Aitman, T. J. *et al.* Copy number polymorphism in Fcgr3 predisposes to glomerulonephritis in rats and humans. *Nature* **439,** 851–855 (2006).
28. Hollox, E. J. *et al.* Psoriasis is associated with increased β-defensin genomic copy number. *Nature Genet.* **40,** 23–25 (2008).
29. Gonzalez, E. *et al.* The influence of *CCL3L1* gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* **307,** 1434–1440 (2005).
30. Bailey, J. A. *et al.* Recent segmental duplications in the human genome. *Science* **297,** 1003–1007 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank H. Mefford, A. Itsara, G. Cooper, T. Brown and G. McVicker for comments during the preparation of this manuscript. The authors are also grateful to J. Sikela and L. Dumas for assistance with the comparison to cDNA microarray data sets. We are grateful to L. Faust, J. Rogers, Southwest National Primate Research Center (P51-RR013986) and P. Parham for providing some of the primate material used in this study and to M. Adams for providing the alignments for the positive selection analysis. We also thank the large genome sequencing centres for early access to the whole genome sequence data for targeted analysis of segmental duplications. This work was supported, in part, by an NIH grant HG002385 to E.E.E. and NIH grant U54 HG003079 to R.K.W. and E.R.M. INB is a platform of Genoma España. T.M.-B. is supported by a Marie Curie fellowship and by Departament d'Educació i Universitats de la Generalitat de Catalunya. E.E.E. is an investigator of the Howard Hughes Medical Institute.

Author Contributions E.E.E. planned the project. M.V. and M.F.C. performed the FISH experiments. T.A.G., L.W.H., L.A.F., E.R.M. and R.K.W. generated the orang-utan WGS sequences. T.M.-B., J.M.K., Z.C., Z.J., L.C., E.E.E. and S.G. analysed the data. C.B. performed the array-CGH experiments. T.M.-B., R.M.-B. and P.S. characterized the chromosome 10 expansion. C.A. and G.A. generated the Venter/Watson comparative duplication maps. A.N. developed the maximum likelihood evolutionary model. T.M.-B., J.M.K. and E.E.E. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to E.E.E. (eee@gs.washington.edu).

Vol 457|12 February 2009|doi:10.1038/nature07684

LETTERS

Human occludin is a hepatitis C virus entry factor required for infection of mouse cells

Alexander Ploss[1]*, Matthew J. Evans[1]†*, Valeriya A. Gaysinskaya[1], Maryline Panis[1], Hana You[1], Ype P. de Jong[1,2] & Charles M. Rice[1]

Hepatitis C virus (HCV) is a leading cause of liver disease worldwide. The development of much needed specific antiviral therapies and an effective vaccine has been hampered by the lack of a convenient small animal model. The determinants restricting HCV tropism to human and chimpanzee hosts are unknown. Replication of the viral RNA has been demonstrated in mouse cells[1,2], but these cells are not infectable with either lentiviral particles bearing HCV glycoproteins (HCVpp)[3] or HCV produced in cell culture (HCVcc) (A.P., M.E. and C.M.R., unpublished observations), suggesting that there is a block at the level of entry. Here we show, using an iterative complementary DNA library screening approach, that human occludin (OCLN) is an essential HCV cell entry factor that is able to render murine cells infectable with HCVpp. Similarly, OCLN is required for the HCV-susceptibility of human cells, because its overexpression in uninfectable cells specifically enhanced HCVpp uptake, whereas its silencing in permissive cells impaired both HCVpp and HCVcc infection. In addition to OCLN, HCVpp infection of murine cells required expression of the previously identified HCV entry factors CD81 (ref. 4), scavenger receptor class B type I (SR-BI, also known as SCARB1)[5] and claudin-1 (CLDN1)[6]. Although the mouse versions of SR-BI and CLDN1 function at least as well as the human proteins in promoting HCV entry, both OCLN and CD81 must be of human origin to allow efficient infection. The species-specific determinants of OCLN were mapped to its second extracellular loop. The identification of OCLN as a new HCV entry factor further highlights the importance of the tight junction complex in the viral entry process, and provides an important advance towards efforts to develop small animal models for HCV.

HCV virions, lipid-enveloped nucleocapsids bearing the viral glycoproteins E1 and E2, seem to enter a host cell in a highly coordinated process involving components of the virus particle and numerous cellular factors[7]. From the long list of putative HCV entry factors, strong evidence supports specific roles for the tetraspanin CD81 (ref. [4]), SR-BI[5] and the tight junction protein CLDN1 (ref. [6]). This list appears incomplete, however, as numerous human cell lines and all non-primate cell lines are resistant to HCV entry even when all three human factors are overexpressed (refs [3,6] and data not shown).

To identify additional factors that can render non-human cells susceptible to HCV entry, we performed a cyclic retrovirus-based repackaging screen of a cDNA library derived from the highly HCV-permissive human hepatocarcinoma Huh-7.5 cell line[8]. We screened for genes that rendered a non-permissive mouse embryonic fibroblast cell line, NIH3T3, infectable with HCVpp, defective lentiviral particles that display the HCV glycoproteins and measure only viral entry[9–11] (Supplementary Fig. 1a and Supplementary Methods). To maximize the likelihood of identifying a new component that would permit HCV entry into non-human cells, we used an NIH3T3 subclone overexpressing human CD81, SR-BI and CLDN1. This screen identified human OCLN as a potential HCV entry factor. OCLN is a four transmembrane domain protein present in the tight junction complex of polarized epithelial cells, where it probably functions to regulate paracellular permeability and cell adhesion[12,13]. Expression of human CD81, SR-BI, CLDN1 and OCLN in NIH3T3 cells enhanced HCVpp infectivity by approximately 120-fold (Fig. 1c and Supplementary Fig. 1c).

To confirm that OCLN is required for HCV entry, we examined the relationship between its expression and HCV susceptibility in a variety of human cell lines. Naturally HCV-permissive human hepatoma cell lines (Huh-7.5 and Hep3B) or those previously shown to lack other entry factors (hepatocellular carcinoma HepG2 and embryonic kidney 293T cells) were found to express readily detectable levels of endogenous OCLN (Supplementary Fig. 1d). In these cells, further overexpression of OCLN did not enhance susceptibility to HCVpp (Fig. 1a). Silencing of OCLN, however, inhibited HCVpp infection of Hep3B cells (Fig. 2f), and both HCVpp and HCVcc infection of Huh-7.5 cells (Fig. 2e and g, respectively), indicating that OCLN is essential for HCV infection of naturally permissive cells. This inhibition was HCV-specific, as infection with pseudoparticles bearing the vesicular stomatitis virus glycoprotein (VSVGpp) was not affected (values in Fig. 2e–f are corrected for VSVGpp infectivity). Notably, OCLN knockdown resulted in incomplete inhibition of HCVpp infection for both cell lines. This is probably due to incomplete silencing or stable reservoirs of the protein that remained to contribute some entry functions. In contrast, the naturally HCV-resistant renal carcinoma cell line 786-O expresses high levels of the major HCV entry factors CD81, SR-BI and CLDN1, but approximately 17-fold less OCLN than Huh-7.5 cells (Supplementary Fig. 1d). OCLN overexpression in 786-O cells specifically enhanced HCVpp infection by more than 88-fold (Fig. 1a). Furthermore, this OCLN-dependence of HCVpp infection was observed across a panel of diverse HCV genotypes (Fig. 1b). Another HCVpp-resistant cell line—the HeLa cell-derived cervical carcinoma line TZM—was found to lack both endogenous CLDN1 and OCLN (more than 100- and 25-fold less messenger RNA than Huh-7.5 cells, respectively; Supplementary Fig. 1d). Overexpression of these factors together increased HCVpp infectivity of TZM cells by 450-fold (Fig. 1a). These results indicate that OCLN is an essential HCV entry factor.

To examine HCV entry requirements in non-human cells further, we transduced murine NIH3T3 cells with all combinations of one, two, three or all four human CD81, SR-BI, CLDN1 and OCLN proteins.

[1]Center for the Study of Hepatitis C, The Rockefeller University, New York, New York, 10065 USA. [2]Division of Gastroenterology, Mount Sinai School of Medicine, New York, New York, 10029, USA. †Present address: Department of Microbiology, Mount Sinai School of Medicine, New York, New York, 10029, USA.
*These authors contributed equally to this work.

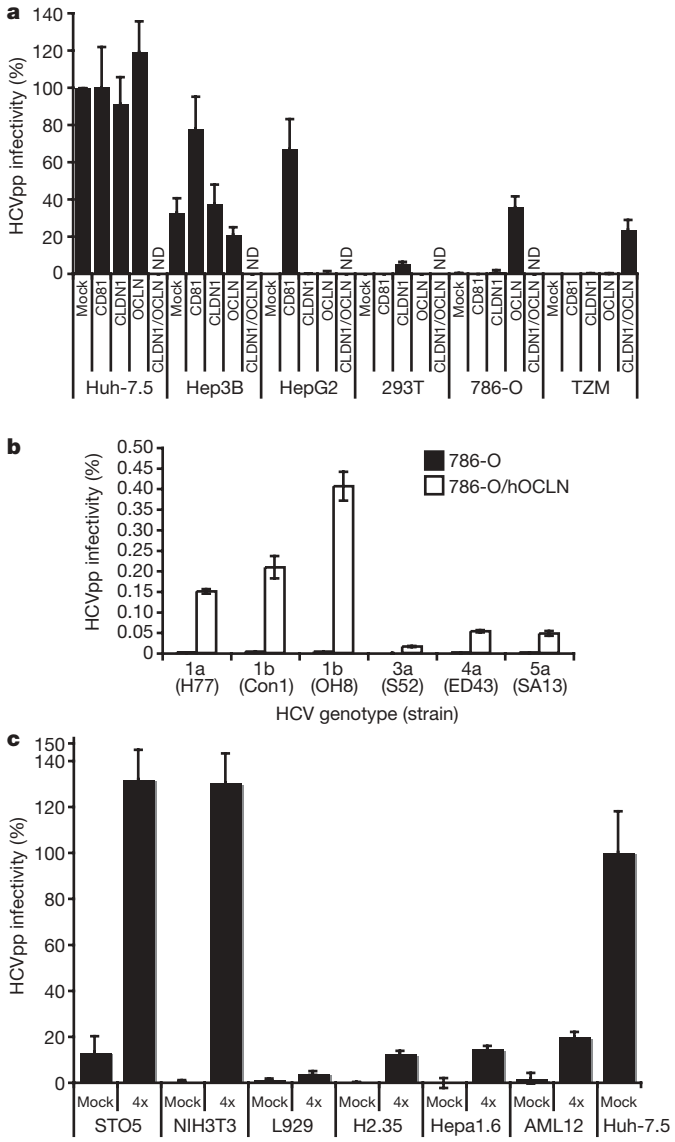


Figure 1 | OCLN expression confers susceptibility to HCVpp. **a**, Human cells were transduced with human pTRIP-mCherry-CD81 (CD81), pTRIP-Cerulean-CLDN1 (CLDN1), pTRIP-Venus-OCLN (OCLN), both CLDN1 and OCLN, or mock transduced, then challenged in parallel with HCVpp and VSVGpp. HCVpp infectivity is reported as the titre of HCVpp divided by the titre of VSVGpp, after subtraction of the signals from infection with non-enveloped pseudoparticles (Env^-pp). HCVpp infectivity is normalized to the signal in Huh-7.5 cells. **b**, Pseudoparticles of different HCV genotypes were used to infect naive (black) or OCLN-expressing (white) 786-O cells. **c**, Mouse cells were transduced with human pTRIP-mCherry-CD81, pTRIP-SR-BI, pTRIP-Cerulean-CLDN1 and pTRIP-Venus-OCLN (4x) or mock transduced, then challenged and the HCVpp infectivity was calculated as above. ND, not determined. Means and s.d. of at least triplicate experiments are shown.

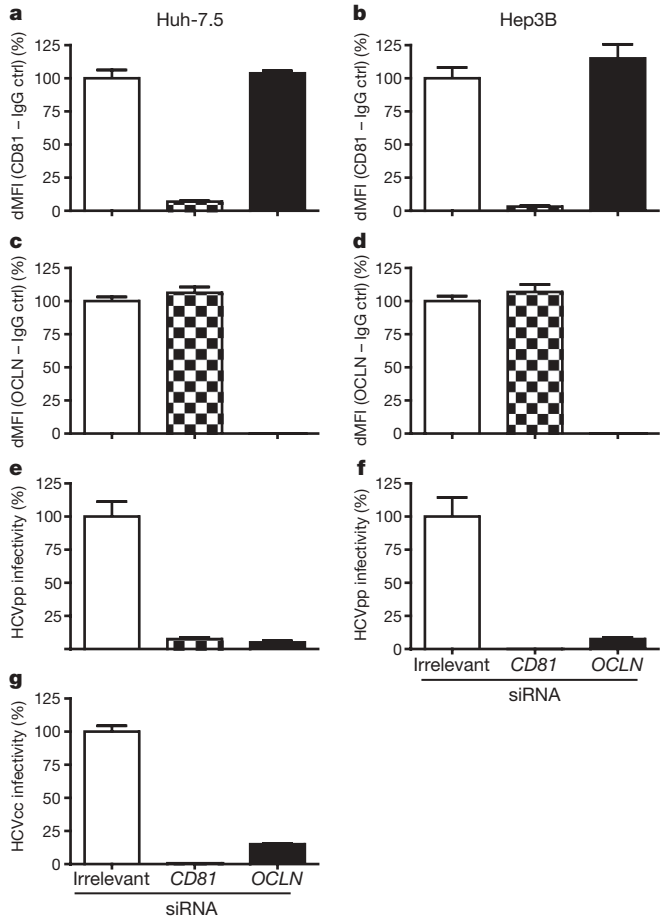


Figure 2 | OCLN silencing inhibits HCV entry. **a**–**g**, Huh-7.5 (left) and Hep3B (right) cells transfected with irrelevant (white), *CD81*-specific (checkered) or *OCLN*-specific (black) short interfering RNA (siRNA) pools were stained with CD81 (**a** and **b**) and OCLN (**c** and **d**) antibodies. Expression was analysed by flow cytometry; the difference between specific and isotype control (IgG ctrl) staining is shown. Concurrently, siRNA-treated cells were challenged with HCVpp (**e** and **f**) and HCVcc (**g**). HCVpp infectivity was calculated as described in methods; HCVcc infection was quantified by the difference between viral protein (NS5A) and isotype control staining. All samples are normalized to irrelevant-siRNA treatment. dMFI, difference in mean fluorescence intensity. Means and s.d. of at least triplicate experiments are shown.

Expression levels were determined by flow cytometry using an mCherry–CLDN1 fluorescent protein fusion (Supplementary Fig. 3a), or specific antibody staining for CD81, SR-BI (Supplementary Fig. 3a) and OCLN (Supplementary Fig. 3b). No combinations of the human factors enhanced VSVGpp infectivity (data not shown and results in Fig. 3a are normalized to VSVGpp values; see Supplementary Fig. 2 for representative raw data). As shown in Fig. 3a, infectivity was not markedly altered by expression of any of the factors alone, or in double or triple combination. Expression of all four entry factors, however, had a large impact on HCVpp permissiveness, with a 45-fold increase over naive NIH3T3 cells. Furthermore, expression of all four factors, expressed as fluorescent protein fusions for CD81, CLDN1 and OCLN and untagged SR-BI, conferred HCV-susceptibility to the mouse embryonic fibroblast cell lines STO5 and L929, as well as the mouse hepatocyte cell lines AML12, Hepa1.6 and H2.35, with enhancements of 4- to 85-fold over naive cells (Fig. 1c). Permissiveness to VSVGpp was not affected by expression of the human proteins. Attempts to infect these murine cells with HCVcc were unsuccessful, probably as a result of inefficient viral RNA replication in these cells, as previously documented[1,2]. It is interesting that different murine cell lines showed varying degrees of HCVpp entry permissiveness after transduction with all four factors. HCV entry factor levels or cell determinants affecting lentivirus reporter expression do not seem to explain the observed differences (because the data were normalized to VSVGpp). Instead, these cell type specific differences may reflect real biological variability in other cellular machinery involved in HCV entry.

To investigate which of the four entry factors are responsible for HCV species-specific tropism, we performed further experiments in Chinese hamster ovary (CHO) cells, which express low endogenous levels of SR-BI, CLDN1 and OCLN (Supplementary Fig. 4b). In CHO cells overexpression of all four human factors specifically enhanced

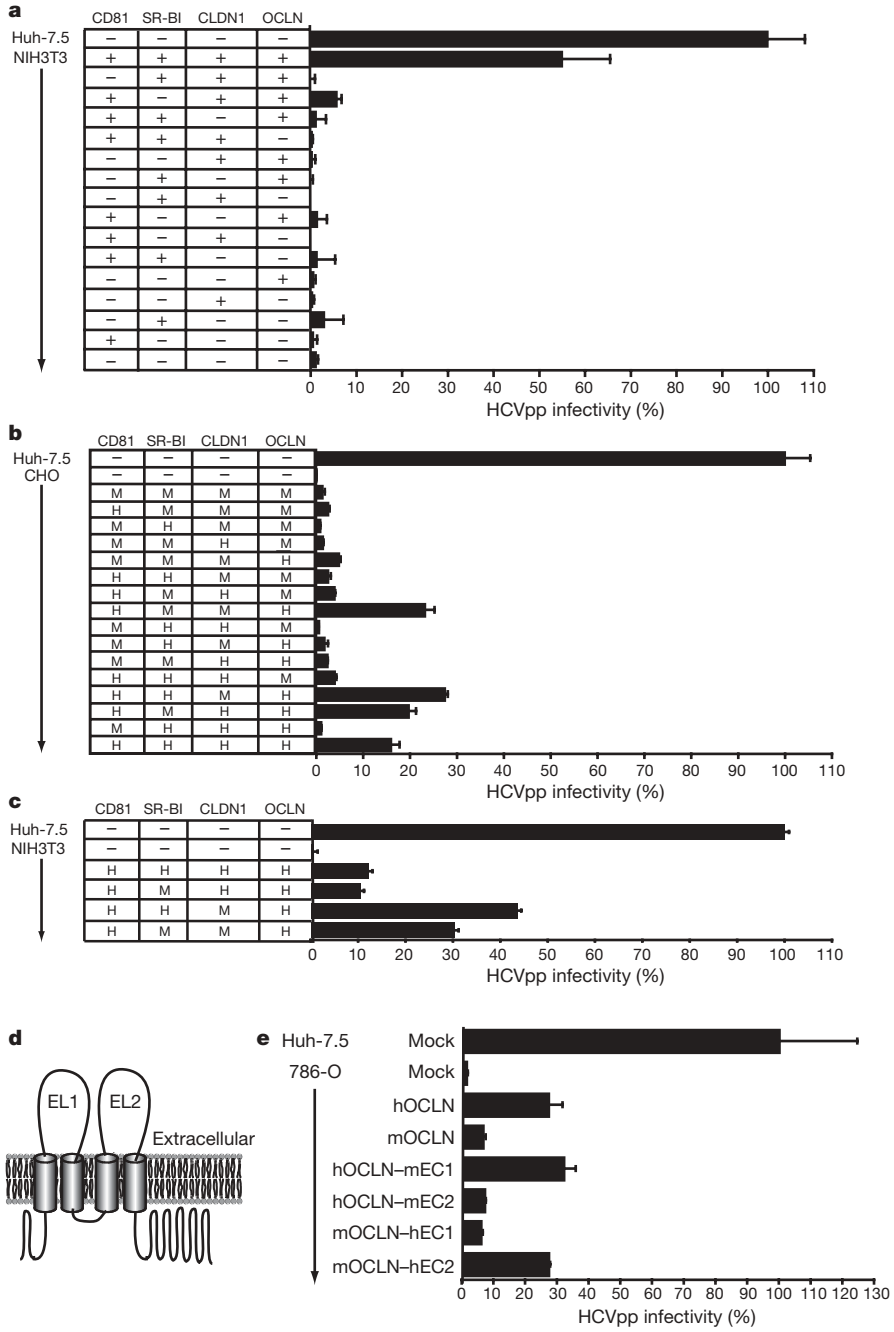


Figure 3 | Expression of human OCLN and CD81 determines HCV species tropism. **a**, Mouse NIH3T3 cells were transduced with combinations of human pTRIP-CD81 (CD81), pTRIP-SR-BI (SR-BI), pTRIP-mCherry-CLDN1 (CLDN1) and pTRIP-OCLN (OCLN). **b**, **c**, Hamster CHO (**b**) and NIH3T3 (**c**) cells were transduced with combinations of human (H) and mouse (M) pTRIP-mCherry-CD81 (CD81), pTRIP-SR-BI (SR-BI), pTRIP-Cerulean-CLDN1 (CLDN1) and pTRIP-Venus-OCLN (OCLN). **d**, Diagram of OCLN membrane topology[12]. **e**, 786-O cells were transduced with mouse or human OCLN chimaeras; all chimaeras were amino-terminally tagged with Venus yellow fluorescent protein. Transduced cells (**a**, **b**, **c** and **e**) were challenged with HCVpp and VSVGpp encoding GFP reporters and HCVpp infectivity calculated and normalized as described in Methods. Means and s.d. of at least duplicate experiments are shown.

HCVpp infection by 60- to 336-fold (Fig. 3b and Supplementary Fig. 4a). Although omission of CD81, CLDN1 or OCLN abolished HCVpp entry, a low level of endogenous SR-BI in CHO cells provided some HCVpp entry function when the other three factors were transduced (Supplementary Fig. 4a). However, because the overexpression of SR-BI further enhanced entry by threefold in the context of the other human proteins (Supplementary Fig. 4a), we still considered CHO cells to be a suitable environment in which to study the species-specific functions of each factor. To do so, we transduced CHO cells to express every combination of human and mouse CD81, SR-BI, CLDN1 and OCLN. In all combinations, mouse SR-BI functioned equivalently to that of human origin (Fig. 3b). This shows for the first time that mouse and human SR-BI are equally capable of mediating HCV uptake. In CHO cells, mouse CLDN1 seems to be slightly more functional for HCVpp entry than the human protein (Fig. 3b). This is in agreement with a slight preference observed when either of these CLDN1 proteins, which are 90% identical, were expressed in human 293T cells[6]. These observations were confirmed by transducing mouse NIH3T3 cells with human CD81 and OCLN and all permutations of mouse and human SR-BI and CLDN1 (Fig. 3c). Although SR-BI showed little to no species specificity, NIH3T3 cells expressing mouse CLDN1, in addition to the other

three entry factors, were slightly more permissive to HCVpp than the same cells expressing human CLDN1. The determinants for this CLDN1 preference have not yet been examined, but it remains clear that CLDN1 does not contribute to the restriction to HCVpp entry in mouse cells. Notably none of the mouse cells examined above express sufficiently high levels of either murine SR-BI or CLDN1 (Supplementary Fig. 4d, e), thus explaining their dependence on transduction of these proteins for HCVpp permissivity.

In contrast, CD81 and OCLN showed human-specific HCV entry factor functions in CHO and NIH3T3 cells. CHO cells expressing either of these human proteins in the context of the other three mouse proteins were slightly more permissive to HCVpp than cells expressing only the mouse proteins (Fig. 3b). Furthermore, expression of both human CD81 and OCLN, together with murine SR-BI and CLDN1, rendered CHO cells as infectable with HCVpp as cells expressing all four human factors (Fig. 3b, compare last bar to any with human CD81 and OCLN). These results agree with our previous finding that mouse and hamster CD81, when expressed in the CD81-deficient human hepatocellular carcinoma cell line HepG2, support only low level HCV entry (greater than tenfold reduced from human CD81)[14]. Importantly, these data indicate that CD81 and OCLN represent the minimal human-specific entry factors, at least in the context of mouse and hamster cells.

For CD81, the species-specific difference in HCV-entry activity between rat and human proteins maps exclusively to its large extracellular loop[14]. To determine the regions of OCLN that are responsible for the functional difference observed between mouse and human proteins, which are 91% identical, we expressed chimaeric OCLN molecules in 786-O cells and assayed HCVpp permissivity. In these experiments, a mouse chimaera bearing the second extracellular loop (EC2) of human OCLN was as active as the full-length human protein and at least fourfold more active than the mouse homologue (Fig. 3e, compare mOCLN–hEC2 to hOCLN and mOCLN, respectively). Conversely, a human OCLN mutant with the EC2 of the mouse protein functioned similarly to the full-length mouse factor in HCVpp infectivity (Fig. 3e, compare hOCLN–mEC2 to mOCLN). These data suggest that the human-specific determinants of OCLN's HCV entry factor functions are entirely contained within EC2.

This study represents an important step forward in understanding both HCV host cell entry and HCV species tropism. All human, murine or hamster cell lines that we tested became permissive for HCV entry when engineered to express the molecules CD81, SR-BI, CLDN1 and OCLN. HCV entry into primary mouse hepatocytes might be complicated by the expression of additional dominant negative restriction factors absent in cell lines, which is experimentally difficult to address at this point. Our data suggest that OCLN completes the list of cell-type specific HCV entry factors; any other factors required for HCV entry must be ubiquitously expressed, or at least conserved between a wide range of human and mouse cells. The fact the OCLN is an important component of the tight junction complex further highlights the significance of this structure and cell polarity to HCV entry. We previously showed that CLDN1 acts late in the entry process, just before virion internalization[6]. The intimate association of CLDN1 and OCLN at the tight junction suggests that both these factors may function in a similar time frame. The use of several uptake factors with distinct cell surface distributions strengthens the hypothesis that HCV follows a coordinated entry pathway similar to that of coxsackievirus B. This virus initially interacts with a primary receptor (decay-accelerating factor) on the luminal cell surface, followed by lateral migration of the virus–receptor complex to the tight junction, where interaction with the coxsackievirus and adenovirus receptor is immediately followed by uptake into the host cell[15]. Notably, coxsackievirus B entry also requires OCLN[16], further suggesting similar entry mechanisms of this virus and HCV. Recent work supports this stepwise model of HCV entry, demonstrating that initial engagement of CD81 on the cell surface, by either fluorescently labelled CD81 antibodies or soluble forms of the HCV glycoproteins, is followed by GTPase-dependent actin rearrangements that allow lateral movement of the CD81-bound complex into areas of cell–cell contact overlapping with both CLDN1 and OCLN localization[17].

The study of HCV pathogenesis and the development of urgently needed effective antivirals and therapeutic and/or preventative vaccines targeting this virus have been severely hampered by the lack of convenient inbred small animal models capable of supporting HCV infection and replication. Numerous blocks will certainly need to be overcome before complete viral replication in a mouse can be achieved. HCV RNA replication in mouse cells is inefficient[1,2], and the ability of such cells to support virion assembly is unknown. Our results clearly demonstrate, however, why mouse cells are unable to support HCV entry. This major block to HCV replication in murine cells can now be overcome simply by the expression of human CD81 and OCLN in the context of mouse CLDN1 and SR-BI, providing a clear foundation upon which a mouse model for HCV infection can begin to be constructed.

METHODS SUMMARY

For more detailed methods see Supplementary Methods.

Pseudoparticles. Pseudoparticles were generated by FuGENE6 (Roche Applied Science)-mediated cotransfection of 293T cells with plasmids encoding (1) a minimal HIV (pTRIP, CSGW, CSPW) or MLV (LMN8) provirus encoding green fluorescent protein (GFP) or other transgene; (2) gag-pol (HIV or MLV); and (3) appropriate viral glycoproteins (HCV E1E2 or VSV-G). To generate Env$^-$pp the glycoprotein vector was replaced with empty vector. Pseudoparticle-containing supernatants were collected at 24, 48 and 72 h, pooled and filtered (0.45 μm pore size). Pseudoparticle infections were performed in the presence of 4 μg ml^{-1} polybrene. A minimum of 72 h elapsed between transduction and reporter gene quantification by flow cytometry or subsequent experiments.

Cyclic lentivirus cDNA library repackaging rescue screen. An Huh-7.5-derived cDNA library in the LMN8 retroviral vector (provided by P. Bieniasz) was packaged into VSVGpp and applied to NIH3T3 cells stably expressing human CD81, SR-BI and CLDN1 (N3xF26—NIH3T3 cells expressing 3 factors, FACS, clone 26), which were then challenged with HCVpp, encoding a puromycin (CSPW) resistance gene, and subjected to antibiotic selection (Supplementary Fig. 1a). Surviving clones were pooled and a fraction was tested for their susceptibility to GFP-encoding HCVpp. The populations were subsequently transfected with MLV gag-pol and VSV-G expression plasmids to package the LMN8 genomes with the cDNA inserts contained in the surviving N3xF26 cells into fresh VSVGpp (cyclic packaging rescue, CPR). These were used to transduce naive N3xF26 cells for subsequent rounds of screening (see Supplementary Methods for further details).

Cell culture grown HCV. A plasmid encoding the chimaeric Jc1Flag2(p7-nsGluc2A) reporter genome[18] was XbaI linearized and transcribed using MEGAscript T7 (Ambion). RNA was electroporated into Huh-7.5 cells using an ECM 830 electroporator (BTX Genetronics) and HCVcc was collected from supernatants 48–72 h after transfection[19].

Received 3 September; accepted 5 December 2008.
Published online 28 January 2009.

1. Uprichard, S. L., Chung, J., Chisari, F. V. & Wakita, T. Replication of a hepatitis C virus replicon clone in mouse cells. *Virol. J.* **3,** 89–97 (2006).
2. Zhu, Q., Guo, J. T. & Seeger, C. Replication of hepatitis C virus subgenomes in nonhepatic epithelial and mouse hepatoma cells. *J. Virol.* **77,** 9204–9210 (2003).
3. Bartosch, B., Dubuisson, J. & Cosset, F. L. Infectious hepatitis C virus pseudo-particles containing functional E1–E2 envelope protein complexes. *J. Exp. Med.* **197,** 633–642 (2003).
4. Pileri, P. *et al.* Binding of hepatitis C virus to CD81. *Science* **282,** 938–941 (1998).
5. Scarselli, E. *et al.* The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *EMBO J.* **21,** 5017–5025 (2002).
6. Evans, M. J. *et al.* Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature* **446,** 801–805 (2007).
7. von Hahn, T. & Rice, C. M. Hepatitis C virus entry. *J. Biol. Chem.* **283,** 3689–3693 (2008).
8. Blight, K. J., McKeating, J. A. & Rice, C. M. Highly permissive cell lines for hepatitis C virus genomic and subgenomic RNA replication. *J. Virol.* **76,** 13001–13014 (2002).
9. Bartosch, B. *et al. In vitro* assay for neutralizing antibody to hepatitis C virus: evidence for broadly conserved neutralization epitopes. *Proc. Natl Acad. Sci. USA* **100,** 14199–14204 (2003).
10. Drummer, H. E., Maerz, A. & Poumbourios, P. Cell surface expression of functional hepatitis C virus E1 and E2 glycoproteins. *FEBS Lett.* **546,** 385–390 (2003).

11. Hsu, M. *et al.* Hepatitis C virus glycoproteins mediate pH-dependent cell entry of pseudotyped retroviral particles. *Proc. Natl Acad. Sci. USA* **100,** 7271–7276 (2003).
12. Paris, L., Tonutti, L., Vannini, C. & Bazzoni, G. Structural organization of the tight junctions. *Biochim. Biophys. Acta* **1778,** 646–659 (2008).
13. Chiba, H. *et al.* Transmembrane proteins of tight junctions. *Biochim. Biophys. Acta* **1778,** 588–600 (2008).
14. Flint, M. *et al.* Diverse CD81 proteins support hepatitis C virus infection. *J. Virol.* **80,** 11331–11342 (2006).
15. Coyne, C. B. & Bergelson, J. M. Virus-induced Abl and Fyn kinase signals permit coxsackievirus entry through epithelial tight junctions. *Cell* **124,** 119–131 (2006).
16. Coyne, C. B., Shen, L., Turner, J. R. & Bergelson, J. M. Coxsackievirus entry across epithelial tight junctions requires occludin and the small GTPases Rab34 and Rab5. *Cell Host Microbe* **2,** 181–192 (2007).
17. Brazzoli, M. *et al.* CD81 is a central regulator of cellular events required for hepatitis C virus infection of human hepatocytes. *J. Virol.* **82,** 8316–8329 (2008).
18. Marukian, S. *et al.* Cell culture-produced hepatitis C virus does not infect peripheral blood mononuclear cells. *Hepatology* **48,** 1843–1850 (2008).
19. Lindenbach, B. D. *et al.* Complete replication of hepatitis C virus in cell culture. *Science* **309,** 623–626 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Bowman, K. Mu, M. Hunter, J. Tassello and M. Holz for excellent technical assistance, J. L. Law and N. Shohdy for advice on siRNA experiments, T. von Hahn and A. J. Syder for helpful discussions and C. Murray for editing the manuscript. S. Mazel, C. Bare and X. Fan provided outstanding technical support. This work was supported by the Greenberg Medical Research Institute, the Ellison Medical Foundation, the Starr Foundation, the Ronald A. Shellow Memorial Fund, the Richard Salomon Family Foundation and funded in part by a Grant from the Foundation for the National Institutes of Health through the Grand Challenges in Global Health initiative (to C.M.R.), and the National Institutes of Health (to M.J.E. and C.M.R.). C.M.R. is an Ellison Medical Foundation Senior Scholar in Global Infectious Diseases. A.P. and M.J.E. were supported by Kimberly Lawrence-Netter Cancer Research Discovery Fund Award Postdoctoral Fellowships. M.J.E. was also supported by the Charles H. Revson Postdoctoral Fellowship. We would like to thank P. D. Bieniasz and T. Hatziioannou for providing reagents, including the LMN8 retroviral plasmid, and expertise necessary for the cDNA library construction and screening. R. Tsien provided the mCherry fluorescent protein encoding plasmid, A. Miyawaki the Venus/YFP plasmid and D. W. Piston the Cerulean/CFP plasmid. Constructs expressing HCV glycoproteins of diverse genotypes and the JFH-1 subgenomic replicon cDNA were provided by J. Bukh and T. Wakita, respectively.

Author Contributions A.P., M.J.E. and C.M.R. designed the project, analysed results and wrote the manuscript. A.P., M.J.E., H.Y., M.P., V.A.G. and Y.P.d.J. performed the experimental work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.M.R. (ricec@rockefeller.edu).

LETTERS

Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter

Michael J. Chen[1,2]†, Tomomasa Yokomizo[3], Brandon M. Zeigler[1], Elaine Dzierzak[3] & Nancy A. Speck[1]†

Haematopoietic stem cells (HSCs) are the founder cells of the adult haematopoietic system, and thus knowledge of the molecular program directing their generation during development is important for regenerative haematopoietic strategies. Runx1 is a pivotal transcription factor required for HSC generation in the vascular regions of the mouse conceptus—the aorta, vitelline and umbilical arteries, yolk sac and placenta[1,2]. It is thought that HSCs emerge from vascular endothelial cells through the formation of intra-arterial clusters[3] and that Runx1 functions during the transition from 'haemogenic endothelium' to HSCs[4,5]. Here we show by conditional deletion that Runx1 activity in vascular-endothelial-cadherin-positive endothelial cells is indeed essential for intra-arterial cluster, haematopoietic progenitor and HSC formation in mice. In contrast, Runx1 is not required in cells expressing Vav1, one of the first pan-haematopoietic genes expressed in HSCs. Collectively these data show that Runx1 function is essential in endothelial cells for haematopoietic progenitor and HSC formation from the vasculature, but its requirement ends once or before Vav is expressed.

HSCs are found in several independent sites in the conceptus, and the transcription factor Runx1, which is required for HSC generation, is expressed in these sites[4]. Runx1 protein is seen in the mesodermal masses and endoderm of the prospective yolk sac blood islands at the neural plate stage[4,6], in the mesoderm of the distal allantois and chorion (precursors of the placenta) at headfold stages[7], and in endothelial cells in the distal allantois, dorsal aorta, and vitelline and umbilical arteries starting the 4–6 somite pair stage (data not shown). Runx1 is expressed at 9.5–10.5 days post coitus (d.p.c.) in both endothelial and mesenchymal cells in the dorsal aorta and placental labyrinth[4,7,8]. Therefore it is not clear whether Runx1 is required in mesoderm, endoderm, mesenchyme and/or endothelial cells for intra-arterial cluster and HSC formation.

To examine the hypothesis that Runx1 is required in endothelial cells *per se*, we deleted *Runx1* in vascular endothelial cadherin (VEC)-positive cells. VEC is a transmembrane protein located at cell-to-cell adherens junctions that is involved in endothelial cell adhesion[9]. VEC is expressed transiently in the yolk sac mesoderm at 7.5 d.p.c.; however, at 8.5 d.p.c. its expression is highly restricted to endothelial cells in the dorsal aorta and heart[10–12]. At 9.5 d.p.c. VEC is expressed in vasculature throughout the conceptus[11]. Distinct mesodermal populations give rise to the yolk sac blood islands and to the major arteries of the conceptus, including those in the placenta[13–15]. Thus, with Cre recombinase expressed from VEC (*Cdh5*) regulatory sequences (VEC-Cre) we could monitor the impact of *Runx1* excision on HSC formation from the vasculature including the dorsal aorta, vitelline and umbilical arteries, and placenta—sites where VEC expression initiates in endothelium.

We generated a *VEC-Cre* transgene (Fig. 1a) and used Cre-reporter mice (*R26R-lacZ* and *R26R-YFP*) to characterize its activity. The first sites of VEC-Cre activity were in the yolk sac and chorionic mesoderm (not shown). Primitive erythrocytes in the yolk sac blood islands were β-galactosidase$^+$ (β-gal$^+$), as were Pecam1$^+$ cells in the vitelline vasculature (Fig. 1b). By 9.5 d.p.c. β-gal activity was high in vasculature throughout the conceptus (Fig. 1c), particularly within Pecam1$^+$ cells in the developing labyrinth of the placenta and in the umbilical artery (Fig. 1d). At 10.5 d.p.c. β-gal$^+$ endothelial cells and intra-arterial clusters were present in the dorsal aorta and vitelline and umbilical arteries, whereas mesenchymal cells were β-gal$^-$ (Fig. 1f). VEC-Cre-mediated excision occurred in 77% (±11%) of cell surface Sca1$^-$ (also known as Ly-6A/E and Ly-6a) VEC$^+$ endothelial cells in the aorta–gonad–mesonephros (AGM) region plus vitelline and umbilical arteries of 11.5-d.p.c. conceptuses and in 86% of Sca1$^+$ VEC$^+$ cells (Fig. 1e; Sca1 marks 50% of HSCs, and VEC marks all HSCs in the AGM region)[2,16]. No excision was detectable in Sca1$^+$ VEC$^-$ cells (Fig. 1e). There was widespread VEC-Cre activity in the fetal liver (Fig. 1f, g), 99% of which was localized to CD45$^+$ (encoded by *Ptprc*), CD71$^+$ (also known as Tfrc$^+$) and/or Ter119$^+$ (Ly76$^+$) blood cells (not shown). Approximately 85% of fetal liver blood cells in 15.5-d.p.c. *R26R-YFP;VEC-Cre* fetuses were YFP$^+$ (Fig. 1h), and are consequently derived from cells that at one time in their history expressed VEC. Ninety-six per cent of CD45$^+$ adult bone marrow cells were YFP$^+$ (Fig. 1i), and thus almost all adult blood is derived from VEC-expressing cells.

VEC-Cre-mediated deletion in mice homozygous for floxed *Runx1* alleles (*Runx1$^{f/f}$*) caused a 65% rate of fetal lethality (Supplementary Table 1), and approximately 10% of 12.5-d.p.c. *Runx1$^{f/f}$;VEC-Cre* fetuses had central nervous system haemorrhaging and fetal liver anaemia characteristic of Runx1 deficiency[17,18] (Fig. 2a and Supplementary Table 1). There was an eightfold decrease in c-kit$^+$ (also known as Kit) cells in the livers of 11.5-d.p.c. *Runx1$^{f/f}$;VEC-Cre* fetuses (Fig. 2b), and in colony-forming unit culture (CFU-C) progenitors in all haematopoietic sites (Fig. 2c). DNA analysis revealed that 84% of colonies derived from *Runx1$^{f/+}$;VEC-Cre* fetal cells had excised the *Runx1^f* allele (Fig. 2d and Supplementary Table 2), consistent with the deletion efficiency determined with the *R26R-YFP* allele. In contrast, only 1 out of 389 colonies picked from cultures established from *Runx1$^{f/f}$;VEC-Cre* fetuses had two deleted *Runx1^f* alleles. Thus, at least one functional *Runx1* allele must be present in a VEC$^+$ cell for CFU-C progenitors to emerge in the fetus.

In vivo repopulation assays revealed that haematopoietic tissues of 11.5-d.p.c. *Runx1$^{f/f}$;VEC-Cre* conceptuses could engraft adult mice, but at a much reduced frequency compared to *Runx1$^{f/+}$;VEC-Cre* cells (Fig. 3a, b). Whereas all donor-derived CFU-C progenitors in the recipients of *Runx1$^{f/+}$;VEC-Cre* HSCs had a deleted *Runx1^f* allele,

[1]Department of Biochemistry, [2]Department of Genetics, Dartmouth Medical School, Hanover, New Hampshire 03755, USA. [3]Department of Cell Biology and Genetics, Erasmus Medical Center, 3000 CA Rotterdam, The Netherlands. †Present address: Abramson Family Cancer Research Institute and Department of Cell and Developmental Biology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA.

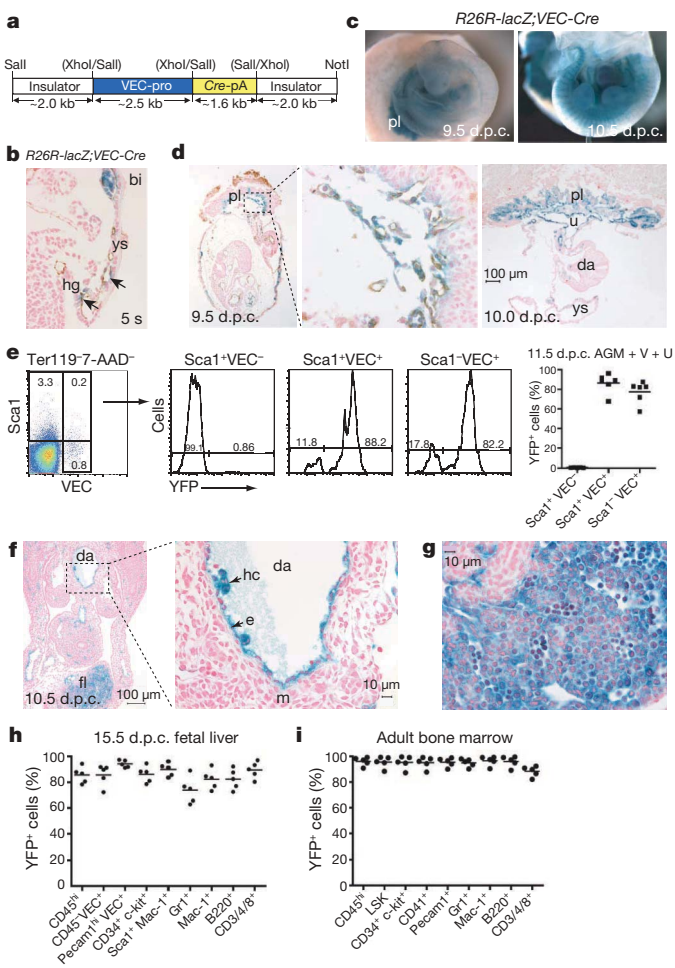


Figure 1 | VEC-Cre excision marks endothelium and blood. **a**, The *VEC-Cre* transgene contains the *Cdh5* promoter (VEC-pro) and *Cre* plus polyA tail (*Cre*-pA), flanked by insulator sequences from the chicken γ-globin gene. **b**, Sagittal section of a 5-somite-pair-stage (5 s) *R26R-lacZ;VEC-Cre* conceptus near the hindgut (hg). Arrows indicate β-gal$^+$ (blue) Pecam1$^+$ (brown) vitelline vasculature. bi, blood island; ys, yolk sac. The blood island contains β-gal$^+$ primitive erythrocytes derived from yolk sac mesoderm in which VEC-Cre was active. **c**, Whole-mount *R26R-lacZ;VEC-Cre* conceptuses; pl, placenta. **d**, VEC-Cre excision in placenta and umbilical artery endothelium. An expanded view of the boxed region in the placental labyrinth (left panel) is shown in the middle panel. Pecam1$^+$ cells are brown. Far right panel shows abundant β-gal$^+$ endothelial cells in the placenta and umbilical artery (u). da, dorsal aorta. **e**, Fluorescence-activated cell sorting (FACS) analysis of cells from the AGM region plus vitelline and umbilical arteries (AGM + V + U) of 11.5-d.p.c. *R26R-YFP;VEC-Cre* conceptuses. Gated populations in the scatter plot were analysed for YFP in histograms to the right. The graph shows percentage of YFP$^+$ cells in each subpopulation; line indicates mean. Means (± s.d.) for each population are: Sca1$^+$ VEC$^-$ (0.7 ± 0.2); Sca1$^+$ VEC$^+$ (86.4 ± 9.9); Sca1$^-$ VEC$^+$ (77.4 ± 11.1). **f**, Left: transverse section through the AGM region of a 10.5-d.p.c. conceptus. Right: expanded image from the boxed region showing β-gal$^+$ endothelial cells (e) and haematopoietic clusters (hc) in the dorsal aorta (da). Note that the sub-aortic mesenchyme (m) is β-gal$^-$. fl, fetal liver. **g**, β-gal$^+$ cells in the liver of *R26R-lacZ;VEC-Cre* fetuses (10.5 d.p.c.). **h**, Percentage of YFP$^+$ fetal liver cells of multiple cell surface phenotypes (gated from Ter119$^-$ 7-AAD$^-$ cells) from 15.5-d.p.c. *R26R-YFP;VEC-Cre* conceptuses. Mean ± s.d. for each population: CD45^{hi} (85.6 ± 6.3); CD45$^-$ VEC$^+$ (85.7 ± 8.2); Pecam1^{hi} VEC$^+$ (94.2 ± 2.8); CD34$^+$ c-kit$^+$ (86.1 ± 6.4); Sca1$^+$ Mac-1$^+$ (89.9 ± 5.1); Gr1$^+$ (also known as Ly6g$^+$; 73.6 ± 10.3); Mac-1$^+$ (82.3 ± 7.6); B220$^+$ (82.4 ± 7.9); CD3/4/8$^+$ (89.4 ± 6.4). **i**, Percentage of YFP$^+$ bone marrow cells (gated from Ter119$^-$ 7-AAD$^-$ cells) in *R26R-YFP;VEC-Cre* adult mice. Mean ± s.d. for each population: CD45^{hi} (96.1 ± 3.3); LSK (95.5 ± 4.4); CD34$^+$ c-kit$^+$ (95.2 ± 4.9); CD41$^+$ (encoded by *Itga2b*; 95.1 ± 4.9); Pecam1$^+$ (95.4 ± 3.5); Gr1$^+$ (95.0 ± 3.2); Mac-1$^+$ (96.5 ± 3.6); B220$^+$ (96.0 ± 4.2); CD3/4/8$^+$ (88.4 ± 4.1).

both *Runx1^f* alleles were deleted in only 1 out of 192 colonies generated from *Runx1$^{f/f}$;VEC-Cre* donor-derived HSCs (Fig. 3c). We also analysed the bone marrow of viable *Runx1$^{f/f}$;VEC-Cre* adult mice and found that all colonies retained at least one undeleted *Runx1^f* allele (Supplementary Table 2). Therefore there is strong selective pressure for cells that escape the excision of at least one *Runx1^f* allele to contribute to haematopoiesis in *Runx1$^{f/f}$;VEC-Cre* mice. Together, these data indicate that Runx1 is required in VEC$^+$ cells for HSC formation.

Because the autonomous production of HSCs occurs first in the major arteries of the conceptus at 10.5 d.p.c. (dorsal aorta as well as vitelline and umbilical arteries)[19], and HSCs localize to the walls of these arteries[2,16], we examined intra-arterial cluster formation in *Runx1$^{f/f}$;VEC-Cre* embryos. VEC is expressed on both endothelial cells and intra-arterial clusters, but the first potential site of deletion is in the endothelium that gives rise to the clusters. No haematopoietic clusters were visible in the dorsal aorta and vitelline and umbilical arteries of *Runx1$^{f/f}$;VEC-Cre* animals, and there was a 90-fold decrease in c-kit^{hi} cells (Fig. 3d). Thus, the important deletion event in these arteries occurred in the endothelium.

To confirm this and to determine whether Runx1 continues to be required once the endothelial-cell-to-HSC transition is complete, we examined the outcome of deleting *Runx1* using Vav-Cre. Vav1 is a GDP/GTP nucleotide-exchange factor for Rho/Rac, the expression of which is restricted to haematopoietic cells[20,21]. *Vav1* was also implicated as a Runx1 target, because its transcript was not detected in the AGM region or fetal liver of Runx1-deficient embryos[22]. Excision of the *R26R-YFP* allele by Vav-Cre occurred in 68% (±16%) of CD45$^+$ VEC$^-$ haematopoietic cells from the AGM region and vitelline and umbilical arteries and not in CD45$^-$ VEC$^+$ endothelium (Fig. 4a). Excised β-gal$^+$ cells were not found in intra-arterial clusters at either 10.5 d.p.c. or 11.5 d.p.c., and were instead located within the sub-aortic mesenchyme and the circulation (Fig. 4b–d). Therefore Vav-Cre marked cells that either originated in the yolk sac or had upregulated Vav-Cre subsequent to their release from the intra-arterial clusters. β-gal$^+$ cells were detectable in the fetal liver at 10.5 d.p.c. (not shown), and by 11.5 d.p.c. 63% (±13%) of CD45$^+$ VEC$^-$ cells in the fetal liver were YFP$^+$ (Fig. 4a) and were easily seen by histology (Fig. 4e). Fetal liver HSCs transiently express cell-surface VEC[23,24]. Sixty-three (±15) per cent of the CD45$^+$ VEC^{lo} HSC-containing population in the 11.5-d.p.c. fetal liver was YFP$^+$ (Fig. 4a), and by 15.5 d.p.c. 90% of lineage-negative Sca1$^+$ Mac-1$^+$ (also known as Itgam$^+$) fetal liver cells were YFP$^+$ (not shown), similar to the excision frequencies reported previously[21]. Thus, with Vav-Cre we could perform an excision in CD45$^+$ cells while avoiding deletion in the intra-arterial clusters or endothelium.

In contrast to VEC-Cre, excision of *Runx1* with Vav-Cre did not affect fetal or adult viability (Supplementary Table 3). The haematopoietic deficiencies described in adult mice upon *Runx1* deletion with Mx1-Cre, including thrombocytopenia, defective lymphopoiesis and expansion of the lineage-negative (Lin$^-$) Sca1$^+$ c-kit$^+$ (LSK) population[25,26], were also apparent in the *Runx1$^{f/f}$;Vav-Cre* mice (Supplementary Fig. 2). Fetal livers from 15.5-d.p.c. *Runx1$^{f/f}$;Vav-Cre* mice had increased numbers of CFU-C progenitors, particularly granulocyte/macrophage (GM) and multipotent (GEMM) progenitors (Fig. 4f), as was previously observed in the bone marrow of Mx1-Cre deleted adult mice[25,26]. Thus, deletion of *Runx1* in Vav$^+$ cells did not impair the emergence of CFU-C progenitors but did affect their homeostasis. Importantly, both *Runx1^f* alleles were deleted in 280 out of 281 colonies from 15.5-d.p.c. *Runx1$^{f/f}$;Vav-Cre* fetal livers (Supplementary Table 4). Colonies from 11.5-d.p.c. animals contained *Runx1^f* alleles at various stages of deletion (Fig. 4g) because excision is initiating close to that time. Nevertheless, 32% of colonies from the haematopoietic tissues of 11.5-d.p.c. fetuses contained two deleted *Runx1^f* alleles, and by 14.5 d.p.c. 100% of fetal liver progenitors had deleted both alleles (Supplementary Table 4). We therefore conclude that Runx1 function is no longer required once a CFU-C progenitor expresses Vav.

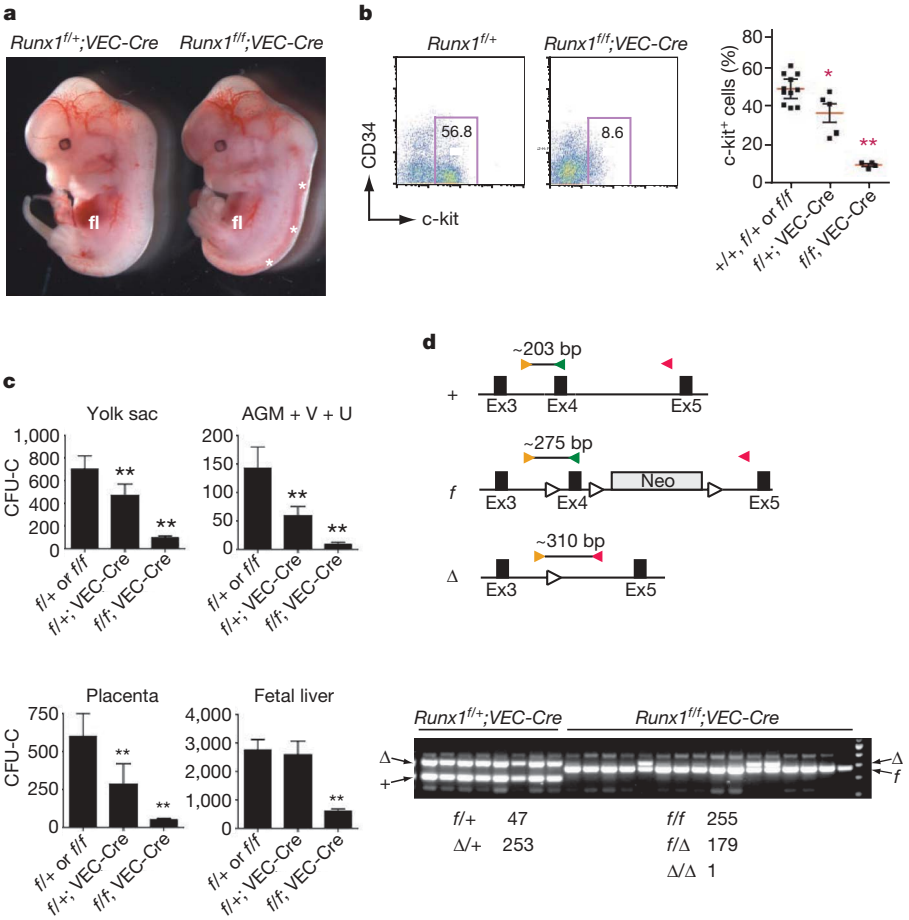


Figure 2 | Runx1 is required in VEC$^+$ cells for haematopoietic progenitor formation. **a**, Gross appearance of *Runx1$^{f/f}$;VEC-Cre* fetuses. Note the central nervous system haemorrhaging (white asterisks) and pale fetal liver (fl). VEC-Cre deletion of *Runx1* resulted in the same phenotype as germline *Runx1* deletion, but the penetrance was lower. **b**, FACS analysis of Ter119$^-$ 7-AAD$^-$ cells from 11.5-d.p.c. fetal livers. Error bars in the plot on right represent 95% confidence intervals. Significant differences from combined +/+, *f*/+ and *f*/*f* samples are indicated with asterisks (*$P = 0.05$, **$P = 0.01$; ANOVA and Dunnett's multiple comparison test). **c**, CFU-C assays of yolk sac, combined AGM region, umbilical and vitelline arteries (AGM + U + V), placenta and fetal liver (represented as numbers per embryo equivalent; error bars indicate 95% confidence intervals). Significant differences from *Runx1$^{f/f}$* or *Runx1$^{f/+}$* conceptuses are indicated with asterisks (**$P = 0.01$). Data are averaged from 16 litters, with 10–31 animals per genotype. **d**, Polymerase chain reaction (PCR) genotyping of single colonies picked from CFU-C assays. The locations of the three primers (coloured triangles) used to detect the wild-type (+), floxed (*f*) and deleted (Δ) *Runx1* alleles are shown at the top. LoxP sites are indicated by open triangles. Neo, neomycin resistance gene. The combined number of colonies from 11.5-d.p.c. AGM + V + U and fetal livers of each genotype are indicated below the gel. Because *Runx1* haploinsufficiency decreases the number of CFU-Cs[18], the percentage of *Runx1$^{f/\Delta}$* colonies may actually underestimate the excision frequency.

Mice were efficiently engrafted with haematopoietic tissues from 11.5-d.p.c. *Runx1$^{f/f}$;Vav-Cre* conceptuses (Fig. 4h), and, in all colonies (94 out of 94) derived from donor *Runx1$^{f/f}$;Vav-Cre* HSCs, both *Runx1^f* alleles were deleted (Supplementary Table 4). The same was true in adult *Runx1$^{f/f}$;Vav-Cre* mice; virtually all CFU-C progenitors had deleted both *Runx1^f* alleles (Supplementary Table 4). Hence, Runx1 is not required in a Vav$^+$ HSC.

In summary, the most essential function of Runx1 is during the transition from VEC$^+$ cells to Vav$^+$ HSCs, although it continues to provide important activity in maintaining normal HSC function and for the terminal differentiation of some lineages[25,26] (Supplementary Fig. 1). Our data demonstrate that the relevant VEC$^+$ HSC precursors are in the haemogenic endothelium. Although the yolk sac mesoderm expresses VEC, in the placenta, vitelline and umbilical arteries, and dorsal aorta, which are independent haematopoietic sites derived from distinct populations of mesoderm[13–15], VEC expression is restricted to intra-arterial clusters and endothelium. Intra-arterial clusters in these sites are absent in *Runx1$^{f/f}$;VEC-Cre* conceptuses,

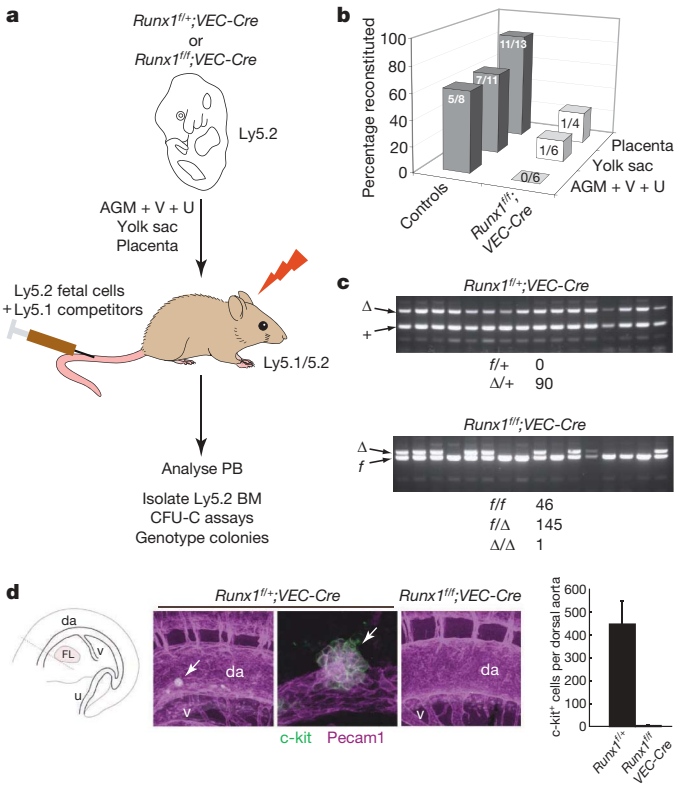


Figure 3 | Runx1 is required in VEC$^+$ cells for HSC emergence. **a**, Scheme of transplantations and analyses. Donor-derived bone marrow (BM) cells were isolated from transplant recipients and plated in methylcellulose assays, and individual colonies were genotyped. **b**, Engraftment as assessed by FACS on peripheral blood (PB) to detect donor-derived (Ly5.1$^-$/5.2$^+$) cells. Numbers indicate successfully reconstituted recipients (>5% donor-derived cells) per number transplanted. Mice were transplanted with one embryo equivalent of 11.5-d.p.c. fetal cells from each site. Control samples are from *Runx1$^{f/f}$*, *Runx1$^{f/+}$* and *Runx1$^{f/+}$;VEC-Cre* conceptuses. **c**, PCR genotyping of single colonies picked from CFU-C assays of donor-derived placenta cells sorted from the bone marrow of transplant recipients. The number of colonies of each genotype is indicated below the gel. **d**, Confocal images of the caudal region of 10.5-d.p.c. embryos (35–36 s) stained with antibodies to Pecam1 (pink) and c-kit (green). The diagram on the left illustrates area analysed (da, dorsal aorta; FL, fetal liver; u, umbilical artery; v, vitelline artery). Arrow indicates a c-kit$^+$ cluster in the ventral aspect of the dorsal aorta. On the right is a bar graph representing the number of c-kit$^+$ cells in the dorsal aorta from three embryos of each genotype. Error bars, standard deviation.

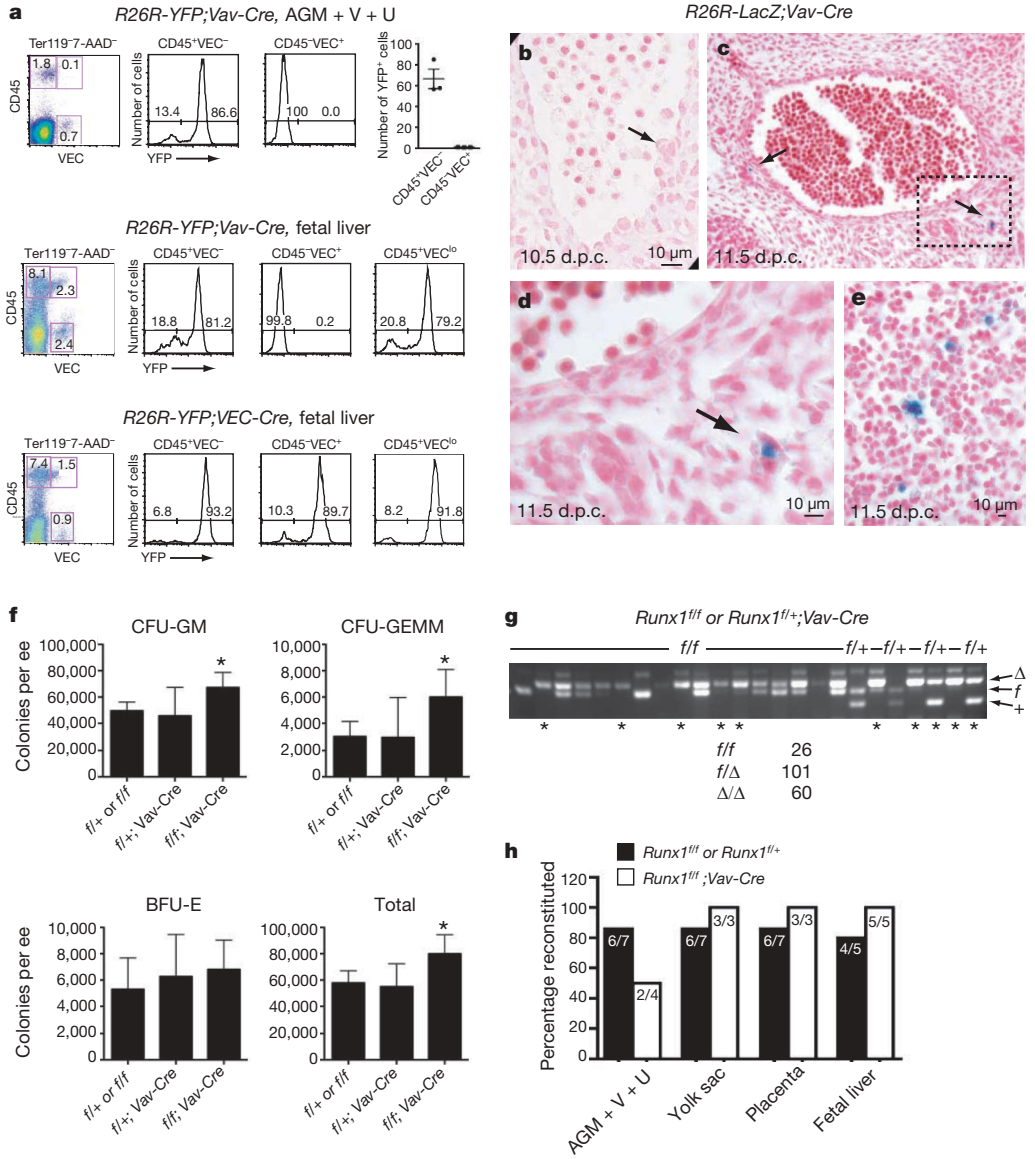


Figure 4 | Runx1 is not required in Vav$^+$ cells for CFU-C or HSC emergence. **a**, Flow cytometric assay of AGM region, vitelline and umbilical arteries (AGM + V + U) and fetal livers from 11.5-d.p.c. *R26R-YFP;Vav-Cre* conceptuses demonstrating deletion in blood but not in endothelium. Scatter plot shows Ter119$^-$ 7-AAD$^-$ cells gated and analysed for CD45 and VEC expression. Histograms to the right represent the percentage of YFP-positive cells in each of the gated populations. The CD45$^+$ VEC^{lo} population from the AGM + V + U was too small to analyse for YFP expression. The top-right graph shows mean ± s.e.m. for AGM + V + U. Activation of the *R26R-YFP* allele in both fetal liver endothelium (CD45$^-$ VEC$^+$) and blood (CD45$^+$ VEC$^-$ and CD45$^+$ VEC^{lo}) with VEC-Cre is shown at the bottom for comparison. **b**, Dorsal aorta from the AGM region of a 10.5-d.p.c. *R26R-lacZ;Vav-Cre* conceptus. The arrow points to an intra-aortic cluster, all of which were β-gal$^-$. **c**, Transverse section through the dorsal aorta of an 11.5-d.p.c. *R26R-lacZ;Vav-Cre* conceptus. Arrows indicate two β-gal$^+$ cells in the subaortic mesenchyme. **d**, Detail of β-gal$^+$ cell in the subaortic mesenchyme from the boxed region in **c**. **e**, β-gal$^+$ cells in the fetal liver (11.5 d.p.c.). **f**, CFU-C assays from 15.5-d.p.c. fetal livers. Numbers are per embryo equivalent (ee). Asterisks indicate significant differences from *Runx1$^{f/+}$* or *Runx1$^{f/f}$* animals ($P = 0.05$, ANOVA and Dunnett's multiple comparison test). Data are compiled from 3–8 conceptuses of each genotype. Error bars represent 95% confidence intervals. **g**, PCR genotyping of colonies from CFU-C assays. The gel is from 11.5-d.p.c. AGM + V + U colonies. Horizontal lines above lanes are *Runx1$^{f/f}$;Vav-Cre* colonies, and *Runx1$^{f/+}$;Vav-Cre* colonies are labelled *f*/+. Asterisks indicate colonies in which both *Runx1^f* alleles (or in the case of *Runx1$^{f/+}$;Vav-Cre* colonies, one *Runx1^f* allele) were completely deleted. Numbers below the gel represent total colonies from 11.5-d.p.c. tissues (AGM + V + U, yolk sac, placenta, fetal liver). **h**, Engraftment of 11.5-d.p.c. tissues (1 ee) as assessed by FACS on peripheral blood to detect donor derived (Ly5.1$^-$/5.2$^+$) cells. Numbers within bars indicate successfully reconstituted recipients/number transplanted.

and therefore the relevant deletion event must have occurred in the endothelium from which they are derived. Continued deletion by VEC-Cre in fetal liver HSCs is unlikely because Runx1 is no longer required once HSCs reach the fetal liver and express Vav, yet they all contain undeleted *Runx1^f* alleles.

We also showed that most adult blood (≥95%) is derived from a cell that at one time in its history expressed VEC. Our working model is that most HSCs are born from VEC$^+$ haemogenic endothelium in the conceptus, although a VEC$^+$ HSC precursor in adult bone marrow cannot be ruled out. It has been shown[27] that a small but detectable number of bone marrow cells (<0.4%) could be labelled when a tamoxifen-regulated VEC-Cre was activated in the adult, hence VEC$^+$ adult HSC precursors, if they exist, are rare cells.

Another model for HSC emergence posits that HSCs in the AGM region originate in subaortic patches of cells in the ventral para-aortic mesenchyme of the dorsal aorta[28], and migrate towards the aorta and between endothelial cells to form the intra-aortic clusters. Because we showed that almost all blood is derived from VEC$^+$ cells, subaortic

patches of cells must upregulate VEC as they protrude through the endothelium.

METHODS SUMMARY

Conditional deletion by VEC-Cre or Vav-Cre. A 2.5 kilobase (kb) VEC (*Cdh5*) promoter sequence was cloned between insulators from the chicken γ-globin gene[29] and used to drive Cre expression. The construct was injected into fertilized oocytes from C57BL6/J × 129S1/SvImJ F_1 mice. *R26R-lacZ/+; VEC-Cre* or *R26R-YFP/+; VEC-Cre* conceptuses were generated by crossing *VEC-Cre/+* males with *R26R-lacZ/+* or *R26R-YFP/+* females. *Runx1*$^{f/+}$*; VEC-Cre* and *Runx1*$^{f/f}$*; VEC-Cre* conceptuses were generated by crossing *Runx1*$^{f/+}$*; VEC-Cre/+* males with *Runx1*$^{f/f}$ females. *Runx1*$^{f/+}$*; R26R-YFP/+; VEC-Cre* or *Runx1*$^{f/f}$*; R26R-YFP/+; VEC-Cre* conceptuses were generated by crossing *Runx1*$^{f/+}$*; R26R-YFP/+; VEC-Cre/+* males with *Runx1*$^{f/f}$ females. Similar breeding strategies were used for Vav-Cre deletion.

Microscopic and histological analyses. Conceptuses were suspended in phosphate-buffered saline or in a 1:2 mixture of benzyl alcohol and benzyl benzoate, and visualized with a stereomicroscope. X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside; Sigma) staining was performed as described previously[4]. In some cases the conceptuses were also incubated with rat anti-mouse Pecam1 and ABC reagent at 4 °C[30], and then treated with DAB peroxidase substrate. Conceptuses were embedded in paraffin, sectioned and counterstained with nuclear fast red.

To analyse intra-aortic clusters, embryos were fixed and stained with anti-c-kit and anti-Pecam1-antibodies. A 1:2 mix of benzyl alcohol and benzyl benzoate was used to increase the transparency of tissues. Samples were analysed with a confocal microscope using the multi-track sequential mode. Three-dimensional reconstructions were generated from *Z*-stacks (62–87 serial sections).

Haematopoietic assays. Methylcellulose colony-forming assays were performed as described previously[4], and colonies counted after 7 days. Fetal tissues were transplanted as described previously[16].

Immuno-staining was performed with phycoerytherin-, allophycocyanin- or Alexa-Fluor-647-conjugated antibodies. Stained cells were analysed on a Becton Dickinson FACSCalibur flow cytometer. Dead cells were excluded with 7-amino-actinomycin D (7-AAD).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 16 July; accepted 7 November 2008.
Published online 7 January 2009.

1. Cai, Z. *et al.* Haploinsufficiency of AML1/CBFA2 affects the embryonic generation of mouse hematopoietic stem cells. *Immunity* **13,** 423–431 (2000).
2. North, T. E. *et al.* Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. *Immunity* **16,** 661–672 (2002).
3. Jaffredo, T., Gautier, R., Eichmann, A. & Dieterlen-Lièvre, F. Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. *Development* **125,** 4575–4583 (1998).
4. North, T. E. *et al. Cbfa2* is required for the formation of intra-aortic hematopoietic clusters. *Development* **126,** 2563–2575 (1999).
5. Yokomizo, T. *et al.* Requirement of Runx1/AML1/PEBP2αB for the generation of haematopoietic cells from endothelial cells. *Genes Cells* **6,** 13–23 (2001).
6. Samokhvalov, I. M., Samokhvalova, N. I. & Nishikawa, S. I. Cell tracing shows the contribution of the yolk sac to adult haematopoiesis. *Nature* **446,** 1056–1061 (2007).
7. Zeigler, B. M. *et al.* The allantois and chorion, when isolated before circulation or chorio-allantoic fusion, have hematopoietic potential. *Development* **133,** 4183–4192 (2006).
8. Rhodes, K. E. *et al.* The emergence of hematopoietic stem cells is initiated in the placental vasculature in the absence of circulation. *Cell Stem Cell* **2,** 252–263 (2008).
9. Lampugnani, M. G. *et al.* The molecular organization of endothelial cell to cell junctions: differential association of plakoglobin, β-catenin, and α-catenin with vascular endothelial cadherin (VE-cadherin). *J. Cell Biol.* **129,** 203–217 (1995).
10. Breier, G. *et al.* Molecular cloning and expression of murine vascular endothelial-cadherin in early stage development of cardiovascular system. *Blood* **87,** 630–641 (1996).
11. Drake, C. J. & Fleming, P. A. Vasculogenesis in the day 6.5 to 9.5 mouse embryo. *Blood* **95,** 1671–1679 (2000).
12. Yokomizo, T. *et al.* Characterization of GATA-1^{+} hemangioblastic cells in the mouse embryo. *EMBO J.* **26,** 184–196 (2007).
13. Kinder, S. J. *et al.* The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. *Development* **126,** 4691–4701 (1999).
14. Lawson, K. A., Meneses, J. J. & Pedersen, R. A. Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development* **113,** 891–911 (1991).
15. Downs, K. M., Hellman, E. R., McHugh, J., Barrickman, K. & Inman, K. E. Investigation into a role for the primitive streak in development of the murine allantois. *Development* **131,** 37–55 (2004).
16. de Bruijn, M. *et al.* Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta. *Immunity* **16,** 673–683 (2002).
17. Okuda, T., van Deursen, J., Hiebert, S. W., Grosveld, G. & Downing, J. R. AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell* **84,** 321–330 (1996).
18. Wang, Q. *et al.* Disruption of the *Cbfa2* gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. *Proc. Natl Acad. Sci. USA* **93,** 3444–3449 (1996).
19. de Bruijn, M. F. T. R., Speck, N. A., Peeters, M. C. E. & Dzierzak, E. Definitive hematopoietic stem cells first emerge from the major arterial regions of the mouse embryo. *EMBO J.* **19,** 2465–2474 (2000).
20. Ogilvy, S. *et al.* Promoter elements of vav drive transgene expression *in vivo* throughout the hematopoietic compartment. *Blood* **94,** 1855–1863 (1999).
21. Stadtfeld, M. & Graf, T. Assessing the role of hematopoietic plasticity for endothelial and hepatocyte development by non-invasive lineage tracing. *Development* **132,** 203–213 (2005).
22. Okada, H. *et al. AML1*$^{-/-}$ embryos do not express certain hematopoiesis-related gene transcripts including those of the *PU.1* gene. *Oncogene* **17,** 2287–2293 (1998).
23. Kim, I., Yilmaz, O. H. & Morrison, S. J. CD144 (VE-cadherin) is transiently expressed by fetal liver hematopoietic stem cells. *Blood* **106,** 903–905 (2005).
24. Taoudi, S. *et al.* Progressive divergence of definitive haematopoietic stem cells from the endothelial compartment does not depend on contact with the foetal liver. *Development* **132,** 4179–4191 (2005).
25. Ichikawa, M. *et al.* AML-1 is required for megakaryocytic maturation and lymphocytic differentiation, but not for maintenance of hematopoietic stem cells in adult hematopoiesis. *Nature Med.* **10,** 299–304 (2004).
26. Growney, J. D. *et al.* Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype. *Blood* **106,** 494–504 (2005).
27. Monvoisin, A. *et al.* VE-cadherin-CreERT2 transgenic mouse: a model for inducible recombination in the endothelium. *Dev. Dyn.* **235,** 3413–3422 (2006).
28. Bertrand, J. Y. *et al.* Characterization of purified intraembryonic hematopoietic stem cells as a tool to define their site of origin. *Proc. Natl Acad. Sci. USA* **102,** 134–139 (2005).
29. Chung, J. H., Whiteley, M. & Felsenfeld, G. A 5′ element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell* **74,** 505–514 (1993).
30. Inman, K. E. & Downs, K. M. Brachyury is required for elongation and vasculogenesis in the murine allantois. *Development* **133,** 2947–2959 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The authors thank G. Ward for his assistance with flow, K. Downs for technical advice, T. Graf for the Vav-Cre mice, and P. Huber for the *Cdh5* sequences. This work was supported by R01HL091724 (N.A.S.), R01DK54077 (E.D.) and T32 AI-07519 (B.M.Z.). Core services were supported in part by the Norris Cotton Cancer Center (NIH CA23108) and the Abramson Family Cancer Research Institute.

Author Contributions T.Y. performed the experiments in Fig. 3d. B.M.Z. performed the experiments in Fig. 1b, d. M.J.C. performed all the remaining experiments. E.D. participated in the interpretation of the experiments and writing the manuscript. N.A.S. participated in the design and interpretation of the experiments, wrote the manuscript and made the figures.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to N.A.S. (nancyas@exchange.upenn.edu).

METHODS

Generation of *Tg(VEC-Cre)* mice. We isolated a 2.5-kb *Cdh5* promoter sequence (gift from P. Huber)[31] with SalI and XhoI and inserted it into the XhoI site of a 427 ins2x plasmid (gift from Z. Li) between two 2-kb chromosomal insulators from the chicken γ-globin gene[29]. A 1.6 kb SalI fragment encoding Cre was inserted into the XhoI site. The construct was excised from the vector by SalI and NotI and injected into fertilized oocytes from C57BL6/J × 129S1/SvImJ F_1 mice.

PCR genotyping protocols and primers. DNA samples were prepared by the HotSHOT method. Primers and conditions for PCR are listed in Supplementary Table 5.

Embryo and mouse generation. *Runx1*$^{f/f}$ mice (*Runx1*$^{tm3Spe/tm3Spe}$) were described elsewhere[26]. *R26R-lacZ* [B6;129-*Gt(ROSA)26Sor*^{tm1Sho}/J] (JAX SN: 003504) and *R26R-YFP* [B6.129X1-*Gt(ROSA)26Sor*$^{tm1(EYFP)Cos}$J] (JAX SN: 006148) mice were purchased from the Jackson Laboratories. *R26R-lacZ/+;VEC-Cre* or *R26R-YFP/+;VEC-Cre* conceptuses were generated by crossing *VEC-Cre/+* males with *R26R-lacZ/+* or *R26R-YFP/+* females. *Runx1*$^{f/+}$*;VEC-Cre* and *Runx1*$^{f/f}$*;VEC-Cre* conceptuses were generated by crossing *Runx1*$^{f/+}$*;VEC-Cre/+* males with *Runx1*$^{f/f}$ females. *Runx1*$^{f/+}$*;R26R-YFP/+;VEC-Cre or Runx1*$^{f/f}$*;R26R-YFP/+;VEC-Cre* conceptuses were generated by crossing *Runx1*$^{f/+}$*;R26R-YFP/+; VEC-Cre/+* males with *Runx1*$^{f/f}$ females. Transplant recipients were bred by crossing [129S1/SvImJ] (JAX SN: 002448) females with B6.SJL (CD45.1) [B6.SJL-*Ptprc*^{a} *Pepc*^{b}/BoyJ] (JAX SN: 002014) males.

Analysis of *R26R-YFP/+;VEC-Cre* conceptuses by fluorescence microscopy or FACS revealed two distinct phenotypes. In approximately 80% of the conceptuses YFP was expressed only in endothelium and blood, whereas 20% of the conceptuses showed mosaic or ubiquitous YFP expression. The later phenotype can be explained by transient and stochastic expression of the *VEC-Cre* transgene in early blastomeres, a phenomenon also described[21] (and observed) for Vav-Cre transgenic animals. PCR genotyping was performed on yolk sac and limb bud samples from *Runx1*$^{f/+}$*;VEC-Cre* and *Runx1*$^{f/f}$*;VEC-Cre* conceptuses, and only conceptuses retaining more than 50% floxed alleles were considered as endothelial-specific excisions and further analysed.

Microscopic and histological analyses. Conceptuses were suspended in phosphate-buffered saline or in a 1:2 mixture of benzyl alcohol and benzyl benzoate, and visualized with a Leica MZFIII stereomicroscope equipped with a Plan 1X 0.14 (0.025–0.125) objective lens (Leica). A Colour Mosaic 11.2 camera and Spot Insight 4.0 acquisition software (Diagnostic Instruments) were used to acquire images. X-gal (Sigma) staining was performed as described previously[4]. In some cases the conceptuses were also incubated with rat anti-mouse Pecam1 (Mec 13.3; BD Pharmingen), biotinylated anti-rat antibody (Santa Cruz Biotechnologies) and ABC reagent (Vector Laboratories) overnight at 4 °C[30], and then treated with DAB peroxidase substrate (Vector Laboratories). Conceptuses were embedded in paraffin, sectioned (8 μm), and counterstained with nuclear fast red (Vector Laboratories). Images were generated with an Optiphot-2 microscope equipped with ×4/0.10 and ×40/1.0 objective lens (Nikon). All images were processed with Adobe Photoshop CS3 software (Adobe System).

To analyse intra-aortic clusters, embryos were fixed for 20–30 min in 2% paraformaldehyde/PBS on ice, and dehydrated in 100% methanol. After removing the yolk sac, rostral half and limb buds, embryos were rehydrated, treated with streptavidin/biotin blocking kit (Vector Laboratories), stained with 1 mg ml^{-1} anti-c-kit antibody (2B8, BD Biosciences) in blocking buffer (PBS containing 0.4% TritonX-100 and 1% skimmed milk) overnight at 4 °C, and rinsed. Goat anti-rat IgG conjugated to Alexa647 (Invitrogen) was added at 1/500 dilution. Embryos were rinsed, incubated in 1 mg ml^{-1} biotinylated anti-Pecam1-antibody (MEC 13.3, BD Pharmingen) overnight at 4 °C, and rinsed. Cy3-conjugated streptavidin (Jackson ImmunoResearch) was added at 1/1,000 dilution. For negative-control staining, rostral halves of embryos and isotype immunoglobulin controls (eBioscience) were used. After rinsing, dehydration and treatment with BABB, a 1:2 mix of benzyl alcohol and benzyl benzoate was used to increase the transparency of tissues. Samples were analysed with a confocal microscope (Zeiss LSM 510 Meta, Plan-Neofluar ×20/NA 0.5) by using multi-track sequential mode. The pinhole diameter was set at 1 Airy unit and the steps were 2.3 mm per *Z*-section. Three-dimensional reconstructions were generated from *Z*-stacks (62–87 serial sections) with LSM Image Browser (Zeiss).

Haematopoietic assays. Methylcellulose colony-forming assays were performed as described previously[4] and colonies counted after 7 days. Tissues were prepared and transplanted as described previously[16].

Flow cytometry. Immunostaining was performed with phycoerytherin-conjugated antibodies to Pecam1 (MEC 13.3), CD34 (RAM34), CD45 (30-F11), CD45.1 (A20), B220 (RA3-6B2), Gr1 (RB6-8C5) and Sca1 (D7; all from BD Pharmingen). We used allophycocyanin-conjugated antibodies to c-kit (2B8), Mac-1 (M1/70), CD3e (145-2C11), CD4 (RM4-5), CD8a (53-6.7), CD45.2 (104; all from BD Pharmingen) and Thy-1.2 (also known as Thy1, 53-2.1, eBioscience), and PerCP-Cy5.5 conjugated antibody to Ter119 (TER-119) (eBioscience). The antibody to CD144 (VE-cadherin, also known as cadherin 5, 11D4.1, BD Pharmingen) was conjugated with Alexa Fluor 647 using the Monoclonal Antibody Labelling Kit (Invitrogen). Stained cells were analysed on a Becton Dickinson FACSCalibur flow cytometer. Dead cells were excluded with 7-AAD (Invitrogen).

31. Gory, B. *et al.* The *vascular endothelial-cadherin* promoter directs endothelial-specific expression in transgenic mice. *Blood* **93**, 184–192 (1999).

LETTERS

The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage

Christophe Lancrin[1], Patrycja Sroczynska[1], Catherine Stephenson[1], Terry Allen[2], Valerie Kouskoff[3] & Georges Lacaud[1]

It has been proposed that during embryonic development haematopoietic cells arise from a mesodermal progenitor with both endothelial and haematopoietic potential called the haemangioblast[1,2]. A conflicting theory instead associates the first haematopoietic cells with a phenotypically differentiated endothelial cell that has haematopoietic potential (that is, a haemogenic endothelium)[3–5]. Support for the haemangioblast concept was initially provided by the identification during mouse embryonic stem cell differentiation of a clonal precursor, the blast colony-forming cell (BL-CFC), which gives rise to blast colonies with both endothelial and haematopoietic components[6,7]. Although recent studies have now provided evidence for the presence of this bipotential precursor *in vivo*[8,9], the precise mechanism for generation of haematopoietic cells from the haemangioblast still remains completely unknown. Here we demonstrate that the haemangioblast generates haematopoietic cells through the formation of a haemogenic endothelium intermediate, providing the first direct link between these two precursor populations. The cell population containing the haemogenic endothelium is transiently generated during BL-CFC development. This cell population is also present in gastrulating mouse embryos and generates haematopoietic cells on further culture. At the molecular level, we demonstrate that the transcription factor Tal1 (also known as Scl; ref. 10) is indispensable for the establishment of this haemogenic endothelium population whereas the core binding factor Runx1 (also known as AML1; ref. 11) is critical for generation of definitive haematopoietic cells from haemogenic endothelium. Together our results merge the two a priori conflicting theories on the origin of haematopoietic development into a single linear developmental process.

The haemangioblast can be isolated from differentiated mouse embryonic stem cells based on Kdr (previously known as Flk-1) expression[12] and generates a blast colony containing haematopoietic and endothelial cells after four days of culture[6,7] (Supplementary Fig. 1). To investigate the developmental steps leading to the generation of haematopoietic cells, we followed the development of individual blast colonies by time-lapse photography. We found retrospectively that it was a sequential process divided into two stages: after 36–48 h of culture the haemangioblast gave rise to a tight adherent structure, and then non-adherent round cells appeared and proliferated to generate a mature blast colony (Fig. 1a and Supplementary Videos 1 and 2). We observed by fluorescence-activated cell sorting (FACS) that a high proportion of cells were positive for the endothelial marker Tek (previously known as Tie2; ref. 13) after the first day (Fig. 1b, d). A few cells positive for CD41 (also known as Itga2b), which defines haematopoietic commitment both *in vivo* and *in vitro*[14–16], were also detected at this stage. The percentage of $CD41^+$ cells increased markedly by day 2, and by day 4 most cells were $CD41^+Tie2^-$, with around 30% expressing CD45 (also known as Ptprc) as well (Fig. 1b, c). FACS sorting of the four cell populations defined by Tie2 and CD41 confirmed that both primitive erythroid and definitive haematopoietic potentials were confined to the $CD41^+$ fractions (Supplementary Fig. 2). These analyses demonstrate the dynamic nature of blast colony development starting at day 1 with a subpopulation of cells expressing Tie2 and developing to a cell population almost fully $CD41^+$ three days later. Interestingly, a transient $Tie2^{hi}$c-Kit^+ (c-Kit^+ is also known as Kit^+) cell population was detected after 48 h of blast development (Fig. 1d), concomitant with the appearance of tight adherent structures (Fig. 1a). This $Tie2^{hi}$c-Kit^+ population contained both $CD41^-$ and $CD41^+$ cells (Fig. 1e). These results indicate that the $Tie2^{hi}$c-Kit^+ cell population may represent a transitional population from which the first $CD41^+$ cells may originate. Accordingly, the $Tie2^{hi}$c-Kit^+CD41^- fraction displayed a low but significant haematopoietic potential (Supplementary Fig. 3), indicating that $CD41^+$ cells may be produced from this cell population.

The endothelial nature of the $Tie2^{hi}$c-Kit^+CD41^- cell population was supported by homogeneous staining for endothelial markers such as CD31 (also known as Pecam1; ref. 17), Flk-1 (ref. 18) and MECA32 (also known as Plvap; ref. 19 and Fig. 2a). Additionally,

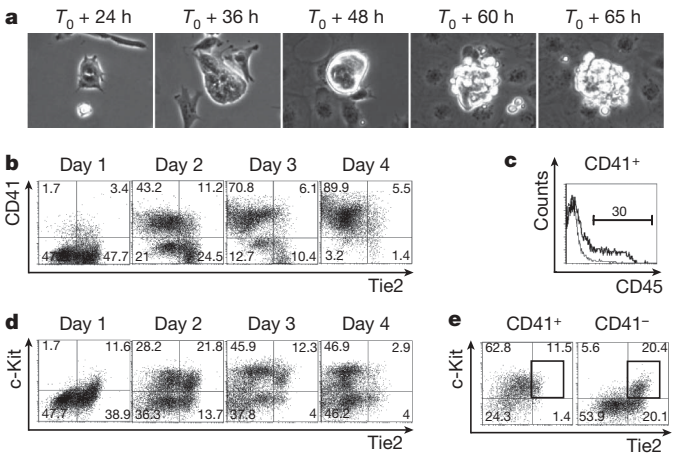


Figure 1 | Analysis of blast colony development. **a**, Phase-contrast time-lapse pictures of blast colony development. **b**, FACS analysis of CD41 and Tie2 expression during blast colony development between days 1 and 4. **c**, CD45 expression of $CD41^+$ cells at day 4 of blast colony culture. The line represents isotype control. **d**, FACS analysis of c-Kit and Tie2 expression during blast colony development. **e**, Tie2 and c-Kit expression of $CD41^+$ (left) and $CD41^-$ (right) cells on day 2 of blast culture. Rectangles indicate the $Tie2^{hi}$c-Kit^+ population. Numbers represent percentages of respective populations.

[1]Cancer Research UK Stem Cell Biology Group, [2]Cancer Research UK Structural Cell Biology Group, [3]Cancer Research UK Stem Cell Haematopoiesis Group, Paterson Institute for Cancer Research, University of Manchester, Wilmslow road, Manchester M20 4BX, UK.

these cells expressed other endothelial genes, such as *Cd34* (ref. 20), endoglin (*Eng*; ref. 21), *Flt1* (ref. 22) and VE-cadherin (*Cdh5*; ref. 23 and Supplementary Fig. 4a), and generated endothelial networks on culture in matrigel plugs (Fig. 2b). Immunostaining for CD31, expressed specifically by Tie2^{hi}c-Kit^{+} cells at day 3 of blast development (Supplementary Fig. 4b), and for CD41 demonstrated the presence of CD31^{+}CD41^{-} endothelial cells corresponding to the core of tightly associated cells of blast colonies (Supplementary Fig. 4c).

In addition to its clear endothelial signature, the Tie2^{hi}c-Kit^{+}CD41^{-} subpopulation expressed transcription factors associated with the onset of haematopoiesis (Supplementary Fig. 5a). To test whether this population contained endothelial cells with haematopoietic potential (that is, a haemogenic endothelium), Tie2^{hi}c-Kit^{+}CD41^{-} cells were cultured in conditions supporting the generation of haematopoietic cells from the aorta–gonad–mesonephros (AGM) region of mouse embryo[24]. After two days, non-adherent round cells were observed (Supplementary Fig. 5b), and FACS analysis demonstrated that around 70% of the cells expressed CD41 and that some of them had down-regulated Tie2 expression (Fig. 2c). Isolation of these newly generated CD41^{+} cells further confirmed that haematopoietic potential was present and restricted to this fraction (Supplementary Fig. 5c). To evaluate the proportion of Tie2^{hi}c-Kit^{+}CD41^{-} cells becoming CD41^{+}, we labelled them with carboxy-fluorescein succinimidyl ester (CFSE), a fluorescent dye equally distributed after each cell division. After 24 h of culture, about 40% of cells were CD41^{+} cells, with a CFSE mean fluorescence 1.5 times lower than in CD41^{-} cells (Supplementary Fig. 5d). This result indicates that CD41^{+} cells were produced from around 25% (that is, 40% divided by 1.5) of the Tie2^{hi}c-Kit^{+}CD41^{-} cells. By limiting dilution analysis, the frequency of Tie2^{hi}c-Kit^{+}CD41^{-} cells generating more differentiated haematopoietic cells was found to be around 1.2% (Supplementary Fig. 5e). This number, lower than in the CFSE analysis, reflects the fact that only a fraction of CD41^{+} cells are able to respond to haematopoietic growth factors. The clonality of the haemogenic endothelium was demonstrated by

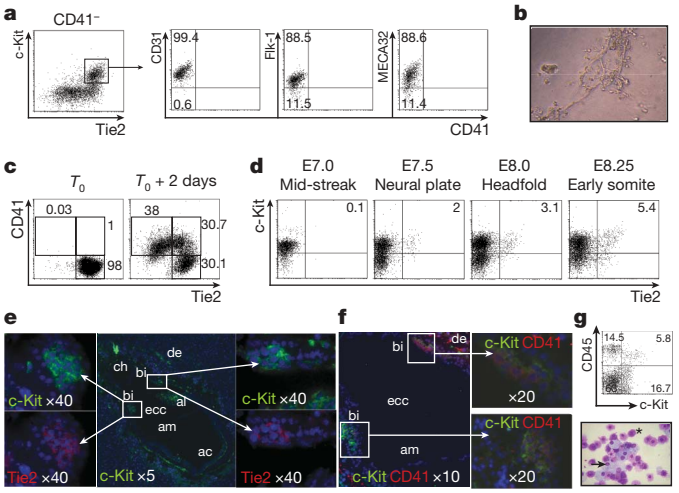


Figure 2 | Tie2^{hi}c-Kit^{+}CD41^{-} cells can generate haematopoietic progenitors. **a**, FACS analyses of CD31, Flk-1 and MECA32 expression by Tie2^{hi}c-Kit^{+}CD41^{-} cells. **b**, Generation of endothelial networks in matrigel by isolated day-2 Tie2^{hi}c-Kit^{+}CD41^{-} cells. **c**, Tie2^{hi}c-Kit^{+}CD41^{-} cells at day 2 of blast development were sorted and cultured. FACS analysis of CD41 and Tie2 expression at T_0 and T_0 + 2 days. **d**, FACS analysis for the presence of Tie2^{+}c-Kit^{+}CD41^{-}cells in gastrulating embryos. **e**, **f**, Immunostaining of gastrulating embryos for Tie2 and c-Kit (**e**) and c-Kit and CD41 (**f**). ac, amniotic cavity; al allantois; am, amnion; bi, yolk sac blood islands; ch, chorion; de, decidua; ecc, exocoelomic cavity. Original magnifications are indicated. **g**, FACS analysis of CD45 and c-Kit expression (top) and May–Grunwald Giemsa staining (bottom) of cells generated by isolated E7.75 mouse embryo Tie2^{+}c-Kit^{+}CD41^{-} cells co-cultured with OP9 cells. Macrophage (arrow) and mast cells (asterisk) are indicated. Numbers indicate percentages of respective populations.

generation of haematopoietic cells from single Tie2^{hi}c-Kit^{+}CD41^{-} cells (Supplementary Fig. 5f). Finally, Tie2^{hi}c-Kit^{+}CD41^{-} cells did not express the mesodermal marker brachyury[25] (Supplementary Fig. 6a) and were unable to generate blast colonies (Supplementary Fig. 6b) indicating that the precursors present in this cell population were clearly distinct and downstream from haemangioblasts. Overall, these data demonstrate the presence of clonal haemogenic endothelial cell precursors in the Tie2^{hi}c-Kit^{+}CD41^{-} cell population.

A prediction from these results is that a Tie2^{+}c-Kit^{+}CD41^{-} haemogenic population should be present *in vivo* at the onset of blood development. We were indeed able to identify Tie2^{+}c-Kit^{+}CD41^{-} cells in embryos at the neural plate stage of gastrulation and their frequencies increased in subsequent stages (Fig. 2d). Immunostainings demonstrated that they were localized in developing blood islands of early head-fold embryos (Fig. 2e, f). FACS-isolated Tie2^{+}c-Kit^{+}CD41^{-} cells plated on OP9 stromal cells generated round non-adherent cells giving rise to CD45^{+} cells and primitive and definitive haematopoietic colonies after replating (data not shown and Fig. 2g). Tie2^{+}c-Kit^{+}CD41^{-} cells were also detected within the AGM region of E10.5 embryos in the dorsal aorta (Supplementary Fig. 7a–c) and these cells also had the capacity to generate haematopoietic cells (Supplementary Fig. 7d, e), but the origin of these Tie2^{+}c-Kit^{+}CD41^{-} cells in the AGM and their precise developmental potential remain to be characterized. Together these data indicate that the Tie2^{+}c-Kit^{+}CD41^{-} progenitor population detected in gastrulating embryos may represent an intermediate between the haemangioblast, predominantly found in the primitive streak[8], and the haematopoietic precursors found in the yolk sac[26].

To investigate further the molecular mechanisms implicated in the generation of the Tie2^{+}c-Kit^{+}CD41^{-} population, we analysed the effects of knockdown of two critical genes for early haematopoiesis, *Runx1* and *Scl*, on this cell population. The transcription factor Runx1 is required *in vivo* for the generation of definitive haematopoietic cells[11] and its absence results *in vitro* in the generation of 20 times fewer blast colonies, with residual blast colonies restricted to a primitive haematopoietic fate[27]. We performed a time-lapse analysis of blast colony development from isolated *Runx1*$^{-/-}$ Flk-1^{+} cells. Clusters of tightly associated adherent cells were observed (Fig. 3a), but only a few of these clusters later generated blast colonies (data not shown). Cells were observed to emerge from most clusters but died instead of proliferating (Fig. 3a and Supplementary Video 3). Immunofluorescence analyses confirmed that cells within these clusters expressed CD31 but that no CD41^{+} cells were present (Supplementary Fig. 8). The defect in haematopoietic development was further observed by FACS analysis that showed a marked reduction in the frequency of CD41^{+} cells (Fig. 3b) and an increased frequency in the Tie2^{hi}c-Kit^{+} cell population (Fig. 3c).

To examine further the requirement for Runx1, we generated a *Runx1*$^{-/-}$ embryonic stem cell line containing a doxycycline-inducible *Runx1* complementary DNA (i*Runx1*). When doxycycline was added at day 2 of blast development, at least a tenfold increase in CD41^{+} cell frequency was observed after 24 h, associated with down-regulation of Tie2 expression (Supplementary Fig. 9a) and induction of expression of genes involved in myeloid cell development (Supplementary Fig. 9b). To establish whether the developmental block observed in the absence of Runx1 was at the level of the Tie2^{hi}c-Kit^{+}CD41^{-} cell population, we isolated these cells and cultured them in the presence of doxycycline. After 48 h, around 60% of the cells expressed CD41 and one-third of these cells down-regulated Tie2 expression (Fig. 3d). Additionally, the CD41^{+} cells incorporated acetylated low-density lipoprotein (LDL), which is taken up by endothelial cells[28], further supporting their endothelial origin (Fig. 3e). Finally, definitive haematopoietic precursors were exclusively detected in colony assays in cultures induced with doxycycline (Fig. 3f). Together, these results demonstrate that the Tie2^{hi}c-Kit^{+}CD41^{-} haemogenic endothelial cell population is generated in absence of Runx1 but that Runx1 is indispensable for the generation of definitive haematopoietic cells from this population.

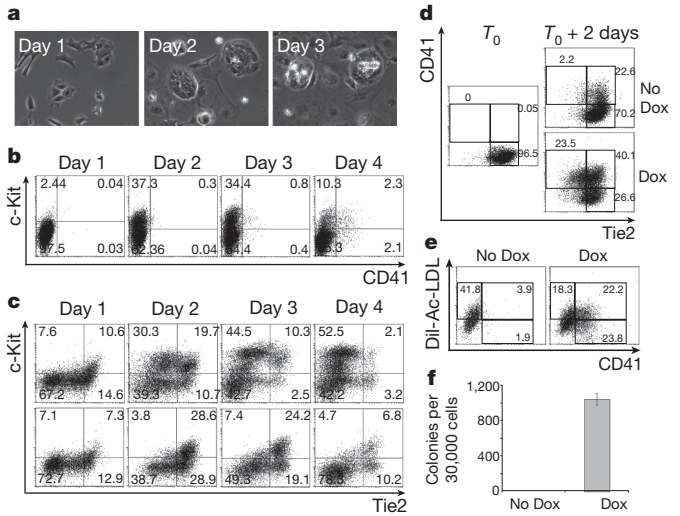


Figure 3 | Runx1 requirement in blast colony development. **a**, Phase-contrast time-lapse photographs of *Runx1*$^{-/-}$ blast colony development. **b**, FACS analysis of CD41 and c-Kit expression during *Runx1*$^{-/-}$ blast colony development between days 1 and 4. **c**, FACS analysis of Tie2 and c-Kit expression during *Runx1*$^{+/+}$ (top) and *Runx1*$^{-/-}$ (bottom) blast colony development. **d**, FACS analysis of Tie2 and CD41 expression after 2 days of culture of i*Runx1* Tie2^{hi}c-Kit^{+}CD41^{-} cells in the absence or presence of doxycycline (Dox; 0.1 μg ml^{-1}). **e**, Dil-Ac-LDL and CD41 expression analysis of i*Runx1* Tie2^{hi}c-Kit^{+}CD41^{-} cells were evaluated in the absence (left) and presence (right) of doxycycline. Numbers indicate the percentage of the respective populations. **f**, Numbers of definitive haematopoietic colonies generated in methylcellulose by i*Runx1* cells collected after two days of culture with or without doxycycline. Error bars indicate standard deviation of the mean ($n = 3$).

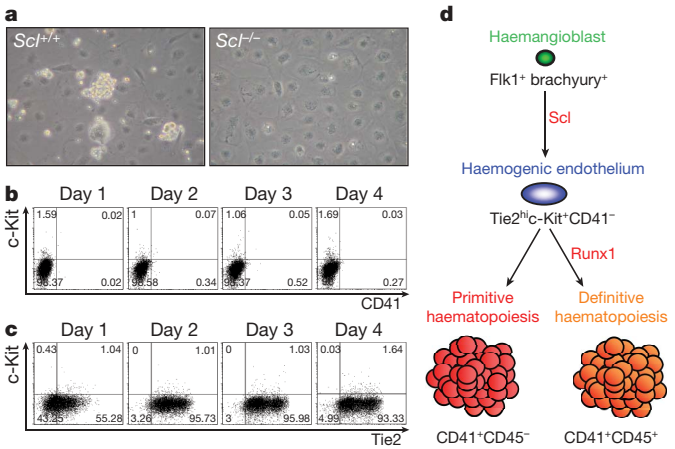


Figure 4 | Scl requirement during blast colony development. **a**, Phase-contrast pictures of day-3 blast colony culture from *Scl*$^{+/+}$ and *Scl*$^{-/-}$ embryonic stem cells. **b**, FACS analysis of CD41 and c-Kit expression during *Scl*$^{-/-}$ blast colony development. Numbers indicate the percentage of the respective populations. **c**, FACS analysis of Tie2 and c-Kit expression during *Scl*$^{-/-}$ blast colony development. **d**, Model of haemangioblast differentiation towards haematopoiesis in the yolk sac.

Accordingly, *Runx1*$^{-/-}$ Flk-1^{+} cells generated significantly more colonies of tightly associated cells than wild-type cells in BL-CFC clonogenic assays (Supplementary Fig. 10). Their number was inversely correlated with the number of blast colonies observed with wild-type cells (Supplementary Fig. 11a). The presence of haemogenic endothelium in these tight structures was supported by detection of a large proportion of CD31^{+} or Tie2^{+}c-Kit^{+} cells in individual colonies by immunostaining (Supplementary Figs 8 and 11b, respectively) and of Tie2^{hi}c-Kit^{+}CD41^{-} cells by FACS analysis on pooled colonies (Supplementary Fig. 11c). Additionally, around 75% of individual colonies (26 out of 34) were able to give rise to haematopoietic cells on Runx1 re-expression (Supplementary Fig. 11d). These data support the notion that these tight colonies generated from the haemangioblast contain cells with haemogenic endothelium potential that are unable to initiate haematopoiesis in the absence of Runx1.

Scl is another critical regulator of haematopoiesis as *Scl*$^{-/-}$ cells are unable to generate either primitive or definitive haematopoiesis[10]. Plating of *Scl*$^{-/-}$ Flk-1^{+} cells resulted, as shown previously[29], in the absence of blast colonies but also in a complete lack of clusters of tightly associated adherent cells (Fig. 4a). We were unable to detect any CD41^{+} cells (Fig. 4b). Although Tie2 was expressed at levels similar to those in *Scl*$^{+/+}$ cells, only a very small fraction of cells co-expressed Tie2 and c-Kit (Fig. 4c) and none of these cells expressed CD31, Flk-1 or MECA32 (data not shown). These data indicate that Scl is critical for the generation of Tie2^{hi}c-Kit^{+}CD41^{-} haemogenic endothelium population and place its role in haematopoietic specification before Runx1 requirement.

In summary, our data provide the first evidence for a direct link between the haemangioblast and a downstream haemogenic endothelium. The generation of this cell population is characterized by the upregulation of c-Kit in a Tie2^{hi} population and this process is Scl-dependent (Fig. 4d). The consecutive generation of definitive haematopoietic cells, characterized by CD41 expression and Tie2 downregulation, requires the transcription factor Runx1, whereas primitive haematopoiesis is Runx1-independent. Identification of these discrete developmental steps will provide the opportunity to explore further the molecular regulation of haematopoietic development.

METHODS SUMMARY

Time-lapse analyses were performed in a Solent Scientific environmental chamber kept at 37 °C for the duration of the analysis and observed under a Zeiss Axiovert 200M microscope. Brachyury–green fluorescent protein (GFP)[25], *Scl*$^{-/-}$ (ref. 10), Ainv18 (ref. 30), Ainv18 *Runx1*$^{+/-}$, *Runx1*$^{-/-}$ and i*Runx1* embryonic stem cell lines were used. Flow cytometry and cell sorting were performed on a FACSCalibur (Becton Dickinson) and FACSVantage or FACSAria cell sorters (Becton Dickinson). For immunostainings, frozen sections were fixed, blocked, incubated with primary Tie2, c-Kit and CD41 antibodies and then with secondary antibodies, mounted and observed under a Zeiss Axiovert 200M microscope.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 8 August; accepted 1 December 2008.
Published online 28 January 2009.

1. Sabin, F. R. Studies on the origin of blood vessels and of red corpuscules as seen in the living blastoderm of the chick during the second day of incubation: contributions to embryology. *Contrib. Embryol.* **9**, 213–262 (1920).
2. Murray, P. D. F. The development *in vitro* of the blood of the early chick embryo. *Proc. R. Soc. Lond.* **11**, 497–521 (1932).
3. Jaffredo, T., Gautier, R., Eichmann, A. & Dieterlen-Lievre, F. Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. *Development* **125**, 4575–4583 (1998).
4. Nishikawa, S. I. *et al.* *In vitro* generation of lymphohematopoietic cells from endothelial cells purified from murine embryos. *Immunity* **8**, 761–769 (1998).
5. North, T. E. *et al.* Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. *Immunity* **16**, 661–672 (2002).
6. Kennedy, M. *et al.* A common precursor for primitive erythropoiesis and definitive haematopoiesis. *Nature* **386**, 488–493 (1997).
7. Choi, K., Kennedy, M., Kazarov, A., Papadimitriou, J. C. & Keller, G. A common precursor for hematopoietic and endothelial cells. *Development* **125**, 725–732 (1998).
8. Huber, T. L., Kouskoff, V., Fehling, H. J., Palis, J. & Keller, G. Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. *Nature* **432**, 625–630 (2004).
9. Vogeli, K. M., Jin, S. W., Martin, G. R. & Stainier, D. Y. A common progenitor for haematopoietic and endothelial lineages in the zebrafish gastrula. *Nature* **443**, 337–339 (2006).
10. Porcher, C. *et al.* The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. *Cell* **86**, 47–57 (1996).
11. Okuda, T., van Deursen, J., Hiebert, S. W., Grosveld, G. & Downing, J. R. AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell* **84**, 321–330 (1996).
12. Faloon, P. *et al.* Basic fibroblast growth factor positively regulates hematopoietic development. *Development* **127**, 1931–1941 (2000).

13. Dumont, D. J., Yamaguchi, T. P., Conlon, R. A., Rossant, J. & Breitman, M. L. *tek*, a novel tyrosine kinase gene located on mouse chromosome 4, is expressed in endothelial cells and their presumptive precursors. *Oncogene* **7,** 1471–1480 (1992).
14. Ferkowicz, M. J. *et al.* CD41 expression defines the onset of primitive and definitive hematopoiesis in the murine embryo. *Development* **130,** 4393–4403 (2003).
15. Mikkola, H. K., Fujiwara, Y., Schlaeger, T. M., Traver, D. & Orkin, S. H. Expression of CD41 marks the initiation of definitive hematopoiesis in the mouse embryo. *Blood* **101,** 508–516 (2003).
16. Li, W., Ferkowicz, M. J., Johnson, S. A., Shelley, W. C. & Yoder, M. C. Endothelial cells in the early murine yolk sac give rise to CD41-expressing hematopoietic cells. *Stem Cells Dev.* **14,** 44–54 (2005).
17. Newman, P. J. *et al.* PECAM-1 (CD31) cloning and relation to adhesion molecules of the immunoglobulin gene superfamily. *Science* **247,** 1219–1222 (1990).
18. Hirai, H. *et al.* Hemogenic and nonhemogenic endothelium can be distinguished by the activity of fetal liver kinase (Flk)-1 promoter/enhancer during mouse embryogenesis. *Blood* **101,** 886–893 (2003).
19. Hallmann, R., Mayer, D. N., Berg, E. L., Broermann, R. & Butcher, E. C. Novel mouse endothelial cell surface marker is suppressed during differentiation of the blood brain barrier. *Dev. Dyn.* **202,** 325–332 (1995).
20. Young, P. E., Baumhueter, S. & Lasky, L. A. The sialomucin CD34 is expressed on hematopoietic cells and blood vessels during murine development. *Blood* **85,** 96–105 (1995).
21. Li, D. Y. *et al.* Defective angiogenesis in mice lacking endoglin. *Science* **284,** 1534–1537 (1999).
22. Fong, G. H., Rossant, J., Gertsenstein, M. & Breitman, M. L. Role of the Flt-1 receptor tyrosine kinase in regulating the assembly of vascular endothelium. *Nature* **376,** 66–70 (1995).
23. Breier, G. *et al.* Molecular cloning and expression of murine vascular endothelial-cadherin in early stage development of cardiovascular system. *Blood* **87,** 630–641 (1996).
24. Mukouyama, Y. *et al.* The AML1 transcription factor functions to develop and maintain hematogenic precursor cells in the embryonic aorta–gonad–mesonephros region. *Dev. Biol.* **220,** 27–36 (2000).
25. Fehling, H. J. *et al.* Tracking mesoderm induction and its specification to the hemangioblast during embryonic stem cell differentiation. *Development* **130,** 4217–4227 (2003).
26. Palis, J., Robertson, S., Kennedy, M., Wall, C. & Keller, G. Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. *Development* **126,** 5073–5084 (1999).
27. Lacaud, G. *et al.* Runx1 is essential for hematopoietic commitment at the hemangioblast stage of development *in vitro*. *Blood* **100,** 458–466 (2002).
28. Voyta, J. C., Via, D. P., Butterfield, C. E. & Zetter, B. R. Identification and isolation of endothelial cells based on their increased uptake of acetylated-low density lipoprotein. *J. Cell Biol.* **99,** 2034–2040 (1984).
29. D'Souza, S. L., Elefanty, A. G. & Keller, G. SCL/Tal-1 is essential for hematopoietic commitment of the hemangioblast but not for its development. *Blood* **105,** 3862–3870 (2005).
30. Kyba, M., Perlingeiro, R. C. & Daley, G. Q. HoxB4 confers definitive lymphoid–myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. *Cell* **109,** 29–37 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. Labib, C. Miller and T. Somervaille for critical reading of the manuscript, J. Barry and M. Hughes for cell sorting, S. Bagley for help with the time-lapse photography, G. Ashton for help with preparation of sections, and L. Gautreau for advice with the OP9 cultures. Cancer Research UK supported this work.

Author Contributions P.S. designed and performed experiments, and analysed the data. C.S. performed experiments. T.A. designed research. C.L., V.K. and G.L. designed the research, performed experiments, analysed the data and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to G.L. (glacaud@picr.man.ac.uk).

METHODS

Embryonic stem cell growth and differentiation. Brachyury–GFP[25], *Scl*$^{-/-}$ (ref. 10), Ainv18 (ref. 30), Ainv18 *Runx1*$^{+/-}$, *Runx1*$^{-/-}$ and i*Runx1* embryonic stem cell lines (data not shown) were used. Growth and differentiation of embryonic stem cells were performed as described previously[25].

Embryo generation. Timed matings of ICR mice were set up and the morning of vaginal plug detection was considered day 0.5. Gastrulating embryos were staged by morphological landmarks. All animal work was performed under regulation in accordance with the UK Animal Scientific Procedures Act (ASPA) 1986.

Adherent BL-CFC and haemogenic endothelium cultures. Flk-1^{+} embryoid body cells were plated on gelatin in the BL-CFC media described previously[25]. For haemogenic endothelium culture, Tie2^{+}c-Kit^{+}CD41^{-} cells were cultured either on gelatin (cells isolated from day-2 or day-3 BL-CFC cultures) or on OP9 stromal cells (cells isolated from embryos) in conditions described previously[24]. Cultures were maintained in a humidified chamber in a normal or low-O_2 (5%) 5% CO_2/air mixture at 37 °C.

Matrigel plug assays. Tie2^{+}c-Kit^{+}CD41^{-} cells isolated from day-2 BL-CFC liquid culture were transferred to 50-μl Matrigel plugs in 96-well plates. Matrigel (BD Biosciences) was diluted 1:1 with IMDM supplemented with 10% FCS, 50 ng ml^{-1} vascular endothelial growth factor and 5 ng ml^{-1} basic fibroblast growth factor and allowed to solidify at 37 °C before the addition of cells. Cultures were maintained at 37 °C, 5% CO_2 and 5% O_2 for five days.

Flow cytometry and cell sorting. Staining was done as described previously[25] and analyses were performed with a FACSCalibur (Becton Dickinson). Sorts were performed with a FACSVantage or a FACSAria (Becton Dickinson). Monoclonal antibodies and streptavidin used were Flk-1–biotin[25], CD31–biotin (MEC13.3, BD Bioscience), MECA32–biotin (BD Bioscience), Tie2–biotin (TEK4, eBioscience), Tie2–phycoerythrin (TEK4, eBioscience), c-Kit–APC (2B8, eBioscience), CD41–FITC (MWReg30, BD Bioscience), streptavidin–phycoerythrin-Cy5 (BD Bioscience) and streptavidin–phycoerythrin-Cy7 (eBioscience). Tie2^{+}c-Kit^{+}CD41^{-} cells from i*Runx1* day-3 blast culture were incubated with Dil-Ac-LDL (10 μg ml^{-1}; Invitrogen Molecular Probes) for 3 h, washed twice to remove all Dil-Ac-LDL, and Runx1 expression was induced with doxycycline. FACS analysis was performed two days later.

CFSE labelling. Tie2^{+}c-Kit^{+}CD41^{-} cells were resuspended in PBS 5% FCS at 0.5×10^6 cells per ml and CFSE was added to 5 μM final concentration. Cells were incubated for 5 min at room temperature (20 °C), washed three times and put back in culture.

Limiting dilution and single-cell analyses. Defined numbers (400, 200, 100 and 50 cells) of Tie2^{hi}c-Kit^{+}CD41^{-} cells were plated on OP9 stromal cells using a FACSAria equipped with an automatic cell deposition unit (Becton Dickinson). For each cell number a minimum of 24 wells was seeded. The cells were first cultured for two days in conditions supporting the transition of haemogenic endothelial cells to haematopoietic progenitors[24] and subsequently in conditions supporting the differentiation to mature haematopoietic cells. The wells were scored for the presence of haematopoietic cells two days later. The haematopoietic nature of the cells was confirmed by May–Grunwald Giemsa staining. The fraction of wells negative for haematopoietic cells was plotted against each dilution and the linear regression was calculated with the Origin software (OriginLab). To obtain the frequency of Tie2^{hi}c-Kit^{+}CD41^{-} cells able to generate haematopoietic cells we used the limdil function of the statistical package R (http://bioinf.wehi.edu.au/software/limdil/index.html). For single-cell analysis, single Tie2^{hi}c-Kit^{+}CD41^{-} cells were sorted on 96-well plates coated with gelatin. The culture conditions were as described previously and the wells were scored for the presence of haematopoietic cells after eight days. The haematopoietic nature of the cells was confirmed by May–Grunwald Giemsa staining.

Immuno-histochemistry. 7–10-μm sections were cut with a Leica CM3050S cryostat from dissected frozen embryos embedded in OCT (Tissue Tek). Sections were fixed with acetone, incubated with 3% H_2O_2, and blocked with 10% goat serum (DAKO) and avidin/biotin blocking kit (Vector). Primary antibodies, Tie2–biotin (TEK4, eBioscience), c-Kit (ACK2, eBioscience) and CD41–biotin (MWReg30, eBioscience) were incubated overnight at 4 °C. Incubations with secondary streptavidin–AlexaFluor555 (Invitrogen) and goat anti-rat IgG2b–HRP (ABD Serotec) were followed by AlexaFluor488 tyramide signal amplification (Invitrogen). The sections were mounted with ProLong Gold antifade reagent with DAPI (4,6-diamidino-2-phenylindole, Invitrogen), analysed using a Zeiss Axiovert 200M microscope and processed with Metamorph software (Universal Imaging).

Time-lapse photography. An Axiovert 200M microscope equipped with a Zeiss ×10 objective lens, a Roper Coolsnap HQ Camera and Solent Scientific environmental chamber kept at 37 °C for the duration of the experiment was used for time-lapse analysis. Phase-contrast images were taken every minute (Supplementary Videos 1 and 3) or every 5 min (Supplementary Video 2) with the image-capture and processing software MetaMorph (Universal Imaging). Video animations were made using the software Imaris (Bitplane). The movies were edited and compressed using Final Cut Studio 2 (Apple).

Gene expression analysis. For gene-specific PCR, total RNA was extracted from each sample with an RNeasy mini kit and treated with RNase-free DNase (Qiagen). 500–1,000 ng of total RNA were reverse-transcribed into complementary DNA with random hexamer using an Omniscript RT kit (Qiagen). The PCR reactions were performed using GoTaq (Promega) and 0.2 μM of each primer. Cycling conditions were as follows: 94 °C for 5 min followed by 30–35 cycles of amplification (94 °C denaturation for 30 s, 60 °C annealing for 30 s, 72 °C elongation for 60 s) with a final incubation at 72 °C for 10 min. The list of primers is available on request.

LETTERS

Continuous single-cell imaging of blood generation from haemogenic endothelium

Hanna M. Eilken[1], Shin-Ichi Nishikawa[2] & Timm Schroeder[1]

Despite decades of research, the identity of the cells generating the first haematopoietic cells in mammalian embryos is unknown[1]. Indeed, whether blood cells arise from mesodermal cells, mesenchymal progenitors, bipotent endothelial–haematopoietic precursors or haemogenic endothelial cells remains controversial[2–9]. Proximity of endothelial and blood cells at sites of embryonic haematopoiesis, as well as their similar gene expression, led to the hypothesis of the endothelium generating blood. However, owing to lacking technology[10] it has been impossible to observe blood cell emergence continuously at the single-cell level, and the postulated existence of haemogenic endothelial cells remains disputed[1]. Here, using new imaging and cell-tracking methods, we show that embryonic endothelial cells can be haemogenic. By continuous long-term single-cell observation of mouse mesodermal cells generating endothelial cell and blood colonies, it was possible to detect haemogenic endothelial cells giving rise to blood cells. Living endothelial and haematopoietic cells were identified by simultaneous detection of morphology and multiple molecular and functional markers. Detachment of nascent blood cells from endothelium is not directly linked to asymmetric cell division, and haemogenic endothelial cells are specified from cells already expressing endothelial markers. These results improve our understanding of the developmental origin of mammalian blood and the potential generation of haematopoietic stem cells from embryonic stem cells.

It has previously been shown that embryonic cells expressing endothelial molecules can generate blood[3,8,11,12]. However, only single markers were used at single time points, and transition of single endothelial cells (ECs) into blood was never continuously observed. We therefore set out to observe the generation of blood cells directly from cells with irrefutable endothelial character, as identified by simultaneous detection of morphological and multiple molecular markers, at the single-cell level, and throughout the whole transition process. To this end, new bio-imaging approaches and well-established differentiation systems of mesodermal cells were combined. Initially, embryonic stem cell (ESC)-derived mesodermal cells[13,14] were used. In numerous studies, this has been confirmed as a valid model of embryonic developmental processes and faithfully reflected *in vivo* development with respect to the sequential appearance of different cell types, expression of developmentally regulated molecules and interdependence of different cell types for inductive signalling[4,8,13–19]. To allow the development of EC colonies before generation of the first blood cells, we chose a well-characterized early mesodermal cell population as the starting population for time-lapse observations. These are purified by sorting fetal liver kinase-1$^+$ (Flk-1, also known as VEGFR-2 and Kdr) epithelial-cadherin$^-$ (E-cadherin, also known as cadherin-1) cells at day 4 of ESC differentiation[20,21] and can differentiate into endothelial, primitive and definitive haematopoietic, cardiomyocytic and mural cells *in vitro*[14]. The generated ECs can integrate into endothelium in embryos and contribute to functional blood vessels *in vivo*[22].

ESC-derived mesodermal precursors were cultured on OP9 stromal cells that support mesodermal differentiation of ESC-derived cells[13,14]. Over the next week, their development was continuously observed at the single-cell level by microscopic time-lapse imaging. With the aid of a computer program, we followed the fate of cells in resulting movies at the single-cell level over days (Fig. 1 and Supplementary Videos 6 and 7). In our time-lapse studies, living blood cells were identified by: non-adherent morphology (Supplementary Fig. 2), proliferation, expression of the blood-specific proteins CD45 (also known as Ptprc), CD41 (Itga2b) and CD11b (Itgam, Mac-1; Supplementary Table 2 and Supplementary Figs 3 and 5), and absence of all following endothelial characteristics. For the identification of ECs, a combination of all relevant criteria currently known for their identification *in vitro* was used: endothelial morphology[13,15,22,23] (Supplementary Fig. 2), expression of vascular endothelial cadherin (VE-cadherin, also known as Cdh5; Supplementary Figs 4 and 5)[24], ability to take up acetylated low-density lipoprotein (Ac-LDL, Supplementary Figs 4 and 5)[25] and presence of functional tight junctions as evidenced by incorporation of claudin 5 at the cell–cell contacts[23,26] (Supplementary Figs 4 and 5). In addition, we only selected EC colonies with sheet morphology[15,23]. Although this ignores EC colonies with less unique morphologies, it is the most stringent approach for EC identification (Supplementary Fig. 2 and Supplementary Table 3).

We found that morphology is a very stringent criterion for the identification of cell and colony types[14,22,23] (Supplementary Table 3) and therefore used it for first experiments. In 25 experiments, the generation of over 4×10^5 cells in more than 10,000 colonies established from individual ESC-derived mesodermal progenitors was continuously observed over up to 7 days. About 7% of all developing mesodermal colonies generated blood cells and 22% of those ($1.5 \pm 0.7\%$ (mean $\pm$ s.d.) of total colonies) had EC sheet morphology (Supplementary Table 1). In 144 EC sheet colonies, we observed adherent cells with clear EC morphology generating non-adherent, living and proliferating cells with haematopoietic morphology (Fig. 1). All haemogenic ECs underwent the same transitions, as witnessed by changes in adherence with highly similar kinetics (Fig. 1b and Supplementary Fig. 7). Haemogenic ECs started to show EC morphology at 5.3 to 6.4 (5.8 ± 0.3) days after the induction of ESC differentiation (Fig. 1b, d). Subsequently, at 5.5 to 6.9 (6.1 ± 0.4) days, the prospective blood cells started to lose tight integration into the EC colony, but still migrated on top of the EC colony in a semi-adherent fashion (Fig. 1e, g). At 6.0 to 8.5 (7 ± 0.6; mean $\pm$ s.d.) days of ESC differentiation, nascent blood cells completely lost adherence to ECs (Fig. 1f, h). Even after observation up to day 11, no later haemogenic transitions were detected. This timing of blood cell emergence is as described in previous reports about blood generation from ESCs[4,13].

[1]Institute of Stem Cell Research, Helmholtz Center Munich—German Research Center for Environmental Health (GmbH), 85764 Neuherberg, Germany. [2]Laboratory for Stem Cell Biology, RIKEN Center for Developmental Biology, 650-0047 Kobe, Japan.

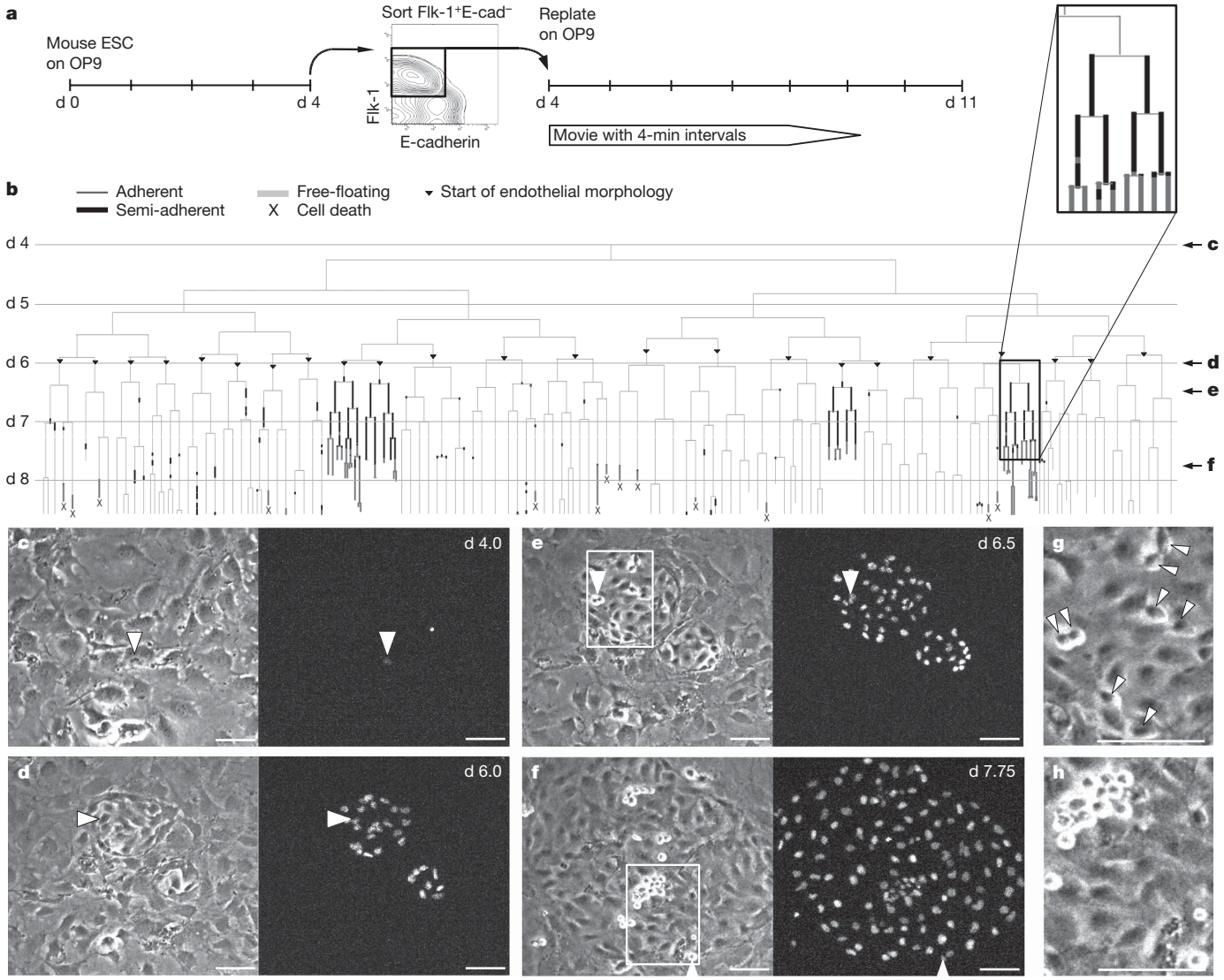


Figure 1 | ESC-derived mesodermal cells give rise to haemogenic endothelial colonies. a, Experimental setup for Figs 1–3. d, day; E-cad, E-cadherin. **b–f**, Single-cell tracking of a haemogenic sheet colony. Cells with endothelial morphology transform into proliferating free-floating cells with haematopoietic morphology (Supplementary Video 1). **b**, Colony pedigree. **c–f**, Phase-contrast (left) and fluorescence images of histone 2B–Venus expression of all ESC (but not OP9) derived cells (right). A single cell (**c**) is forming an endothelial sheet colony (**d**), of which cells (8 arrowheads in **g**) lose integration (**e**, enlarged in **g**) and later completely detach (**f**, **h**). Scale bars, 100 μm.

For increased reliability of EC identification, we established the detection of molecular markers in living cells during time-lapse microscopy. The use of ESCs expressing a fusion of the tight junction protein claudin 5 with the fluorescent protein Venus under the control of regulatory elements of the VE-cadherin gene allowed detection of VE-cadherin expression and EC-specific functional tight junctions[23]. Moreover, the EC-specific ability to take up Ac-LDL was visualized through addition of fluorescently labelled Ac-LDL to the cultures[25]. The haematopoietic nature of nascent blood cells was confirmed by detecting expression of the pan-haematopoietic marker CD45 on living cells using a fluorescently labelled antibody (the same antibody-detection approach was used for additional marker molecules discussed later). Interestingly, only half of Ac-LDL-uptaking and Venus–claudin 5-expressing colonies displayed sheet morphology, whereas all colonies with sheet morphology took up Ac-LDL and expressed Venus–claudin 5 (Supplementary Table 3). This demonstrates that the used combination of morphological and multiple molecular markers for EC identification is more stringent than the use of only single molecular markers in previous studies.

We were able to combine all morphological and molecular markers for endothelial and blood cell identification into a single imaging approach. As shown in Fig. 2, haemogenic ECs first express VE-cadherin, form tight junctions with neighbouring cells, develop endothelial morphology and take up Ac-LDL, but do not express CD45. Subsequently, their ability for Ac-LDL uptake, expression of VE-cadherin, presence of functional tight junctions and adhesion is lost. The cells emigrate from their EC colony, morphologically transform into blood cells, start detectable CD45 expression, and proliferate as suspension cells. Earlier CD45 expression already in VE-cadherin$^+$ cells[11,27] not detected by our live staining approach can not be excluded. All ECs and nascent blood cells express Tie2 (also known as Tek; not shown). The myeloid haematopoietic protein CD11b was co-expressed by most CD45$^+$ cells with a similar onset of expression as CD45 in already free-floating cells (Supplementary Fig. 3 and not shown) proving that a large part of emerging blood cells is not erythroid. CD41 was expressed earlier than CD45 and CD11b, already detectable during the adherent state and overlapping with VE-cadherin::Venus–claudin 5 expression (Fig. 3). This sequence of expression is identical to previous findings showing CD41 to be the first marker for haematopoietic cells emerging *in vivo*[2,28] and from ESCs *in vitro*[29,30]. CD117 (also known as Kit) expression was detectable only on haemogenic ECs and from the semi-adherent stage (not shown), but earlier expression not detected by our live staining

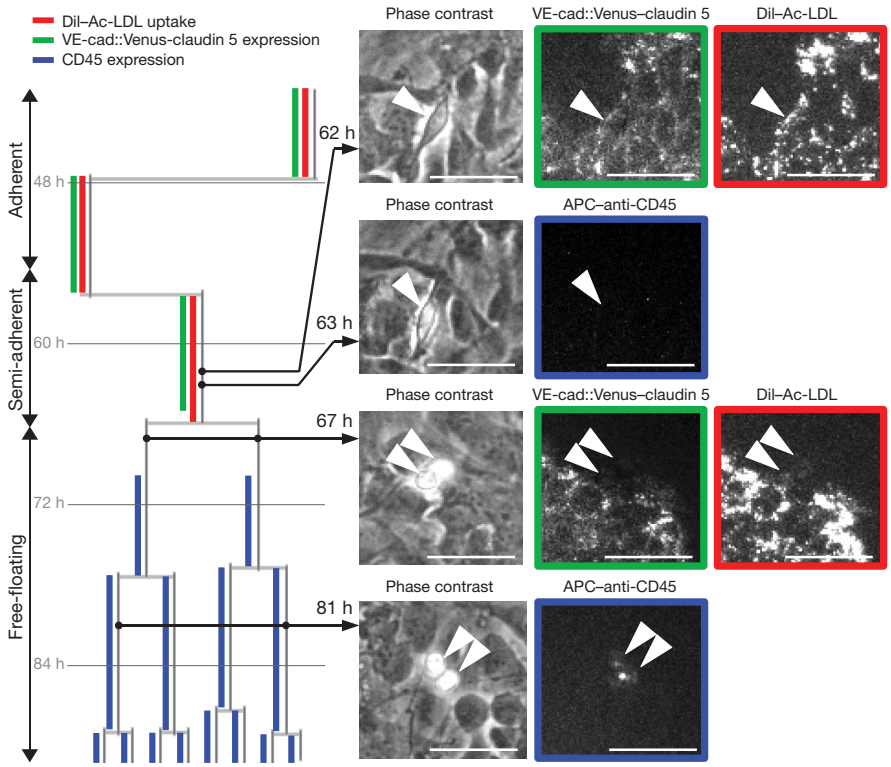


Figure 2 | Continuous observation of endothelial and haematopoietic molecular markers. Part of a haemogenic EC transition (Supplementary Video 2). VE-cadherin::Venus–claudin 5 expression and tight junction formation was detectable at 18 h, and DiI–Ac-LDL uptake at 40 h (not shown). A VE-cadherin-expressing haemogenic EC with tight junctions and DiI–Ac-LDL uptake (62 h) does not express CD45 (63 h). $CD45^+$ cells are visible at the same time point in Supplementary Video 2 as the positive control for CD45 staining. Emerging blood cells subsequently lose endothelial markers and start CD45 expression. Hours after start of the movie and cells of pedigree (arrowheads) are indicated. Scale bars, 50 μm.

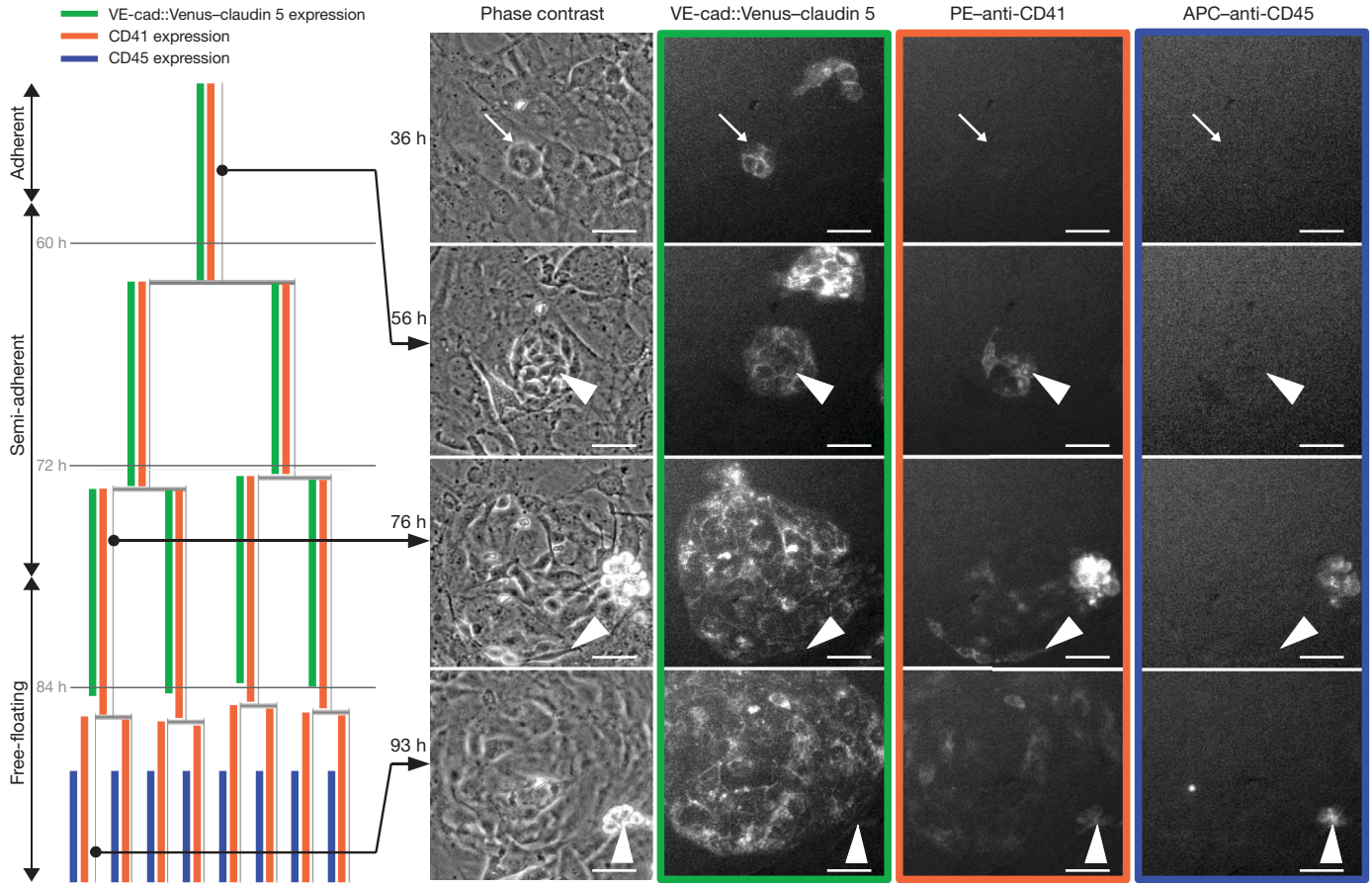


Figure 3 | Nascent blood cells generated from ECs express CD41 before CD45. Part of a haemogenic EC transition (Supplementary Video 3). ECs express VE-cadherin::Venus–claudin 5, but not CD41 or CD45 (36 h, arrow). Some VE-cadherin::Venus–claudin 5^+ cells with sheet EC morphology start expressing CD41 (56 h). $CD41^+$ cells develop into blood cells (76 h) as identified by CD45 staining after becoming free-floating (93 h). Single-cell tracking was started from cells with sheet EC morphology. Hours after start of the movie and cells of pedigree (arrowheads) are indicated. Scale bars, 50 μm.

approach can not be excluded. Extended time-lapse observation to up to a week after the formation of EC colonies showed that ECs can generate haematopoietic progenitor cells with high proliferation potential forming multi-lineage colonies including megakaryocytes (Supplementary Fig. 8 and Supplementary Video 5). Their proliferative potential and the generation of $CD11b^+$ myeloid cells and megakaryocytes indicate that EC-derived blood cells contain definitive haematopoietic progenitors.

The fact that haemogenic ECs are integrated into endothelial sheets by tight junctions strongly supports that they are not early multipotent mesodermal cells expressing EC markers. This was confirmed by demonstrating their inability to generate smooth muscle cells (Supplementary Fig. 6, 12 haemogenic EC colonies, 3 experiments) and functional cardiomyocytes (not shown). The lack of CD45 expression in haemogenic ECs (Fig. 2, 13 haemogenic EC colonies in 2 experiments) rules out that these were bipotent blood and endothelial precursors expressing molecules of both lineages (Fig. 2). All haemogenic ECs fulfilling all described molecular criteria gave rise exclusively to haematopoietic progeny, whereas their endothelial sister cells exclusively generated endothelial progeny (Fig. 1b). Interestingly, their bipotent mother cells already expressed VE-cadherin and incorporated claudin 5 into tight junctions (14 analysed haemogenic EC colonies, 4 experiments) and Tie2 (10 haemogenic EC colonies, 3 experiments). Although Ac-LDL was not incorporated to detectable levels in the short time between onset of other EC markers and haemogenic transformation, this strongly indicates that haemogenic ECs develop from committed ECs that generate both haemogenic and non-haemogenic ECs. In line with *in vivo* observations indicating that only a small subset of ECs may have haemogenic potential[9], we observed that about 1.5% of all ESC-derived mesodermal colonies generate haemogenic sheet ECs (Supplementary Table 1), and that only 1 in about 1,000 ECs is haemogenic at the observed stage. For each haemogenic EC colony, 1–6 haemogenic ECs developed independently of the colony size (4 to over 20 cells in a colony, Fig. 1b and Supplementary Fig. 7). This shows that haemogenic commitment of ECs correlated more with time of culture than with a specific number of mitoses of $Flk\text{-}1^+$ mesodermal cells. Furthermore, the location of haemogenic ECs and their ancestors within their colonies did not influence the haemogenic potential of the ECs. These observations and the highly reproducible timing of haemogenic transitions indicate that a tightly controlled and timed molecular program of intrinsic factors, possibly in combination with extrinsic signals, is essential for the induction of haemogenic commitment in ECs. We did not detect a temporal connection of the loss of adhesion and mitosis of haemogenic ECs. Although not ruling out the possibility that haemogenic commitment could be regulated by asymmetric cell division, this shows that haemogenic loss of adherence and mitosis are not mechanistically linked.

To confirm that ESC-derived haemogenic ECs have an *in vivo* counterpart, we used mesodermal cells derived from mouse embryos. Cells were isolated from 7.5-days-post-coitum embryos[11], and OP9 co-cultures of $Flk\text{-}1^+VE\text{-}cadherin^-CD41^-$ cells were imaged. As for ESC-derived cells, colonies of adherent, Ac-LDL-incorporating primary sheet ECs developed from primary mesodermal cells. Some of these ECs expressed CD41, lost adherence and developed into non-adherent, $CD41^+CD45^+$ haematopoietic cells (Fig. 4 and not shown), demonstrating the generation of haemogenic ECs also from primary cells.

Although our study does not rule out the existence of other haemogenic populations during mammalian embryogenesis, it clearly demonstrates the long-disputed existence of haemogenic endothelium (Supplementary Fig. 1). It shows that haemogenic ECs are transiently generated in a short time window during development and that their

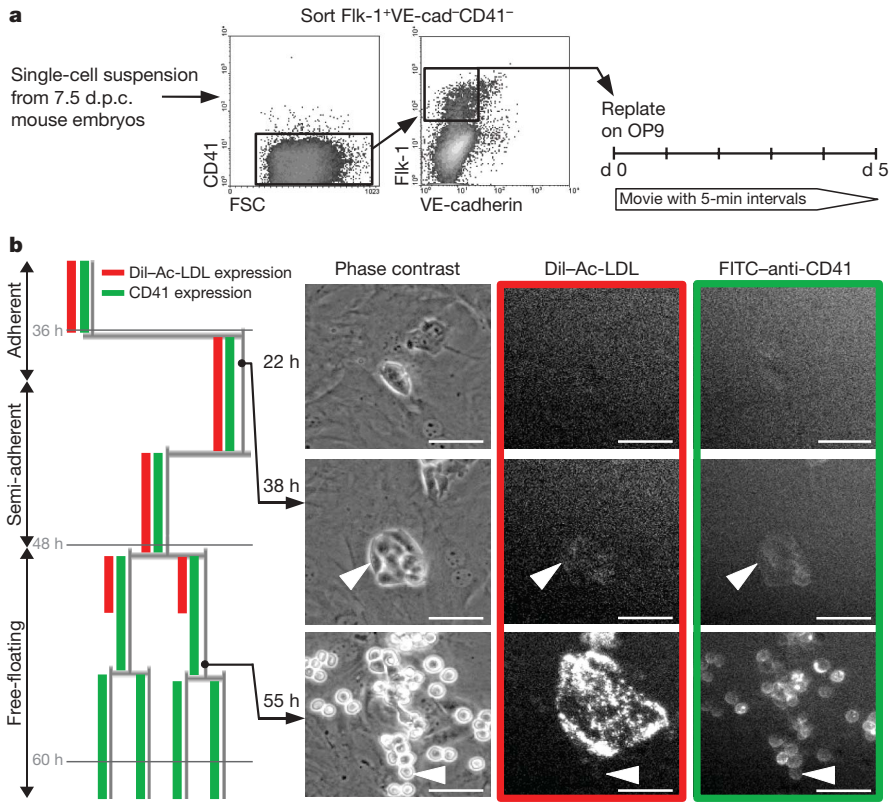


Figure 4 | Mesodermal cells derived from mouse embryos give rise to haemogenic endothelial colonies. Part of a haemogenic EC transition (Supplementary Video 4). Embryonic DiI–Ac-LDL$^-$CD41$^-$ cells generate endothelial colonies taking up DiI–Ac-LDL, in which haemogenic EC also express CD41. In contrast to non-haemogenic ECs, free-floating haematopoietic cells maintain CD41 expression. Single-cell tracking was started from cells with sheet EC morphology. Hours after start of the movie and cells of pedigree (arrowheads) are indicated. Scale bars, 50 μm.

induction probably requires a multitude of intrinsic and potentially extrinsic molecules. By clarifying that ECs can be haemogenic, we lay ground for the future identification of the molecular mechanism leading to the specification of haematopoietic cells in mammalian embryos, and of haematopoietic stem cells from ESCs *in vitro*.

METHODS SUMMARY

Cell culture, isolation of primary cells and FACS sorting were performed as described earlier[11,13]. EB3 and VEVC2 (ref. 23) ESC lines stably expressing histone 2B–fluorescent protein fusions from the pCAG-IRES-Puro plasmid[13] were established as described previously[13].

The following monoclonal antibodies were used: allophycocyanine (APC)–anti-CD45 (clone 30-F11), phycoerythrin (PE)–Cy7 or PE–anti-CD11b (clone M1/70), APC–anti-CD117 (clone 2B8), PE–anti-Tie2 (clone TEK4, all eBioscience), fluorescein isothiocyanate (FITC)– and PE–anti-CD41 (clone MWReg30, BD Pharmingen), biotin–Flk-1 (clone AVAS12), Alexa Fluor 647– or Dylight488–anti-VE-cadherin (clone VECD1), and Alexa Fluor 647–anti-E-cadherin (clone ECCD2). For immunostaining of living cells in movies, antibodies were used at a final concentration of 20 ng ml^{-1} or 100 ng ml^{-1} for CD117.

Time-lapse microscopy for all experiments was performed using a CellObserver system (Zeiss) at 37 °C and 5% CO_2. Time-lapse imaging and single-cell tracking is described in detail in Methods. The risk of confounding identities of adjacent cells during tracking was minimized by use of ESCs constitutively expressing fluorescent proteins with nuclear localization (Fig. 1c–f) and high imaging frequencies (Supplementary Videos 6 and 7).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 29 May 2008; accepted 9 January 2009.

1. Godin, I. & Cumano, A. The hare and the tortoise: an embryonic haematopoietic race. *Nature Rev. Immunol.* **2,** 593–604 (2002).
2. Bertrand, J. Y. *et al.* Characterization of purified intraembryonic hematopoietic stem cells as a tool to define their site of origin. *Proc. Natl Acad. Sci. USA* **102,** 134–139 (2005).
3. Zovein, A. *et al.* Fate tracing reveals the endothelial origin of hematopoietic stem cells. *Cell Stem Cell* **3,** 625–636 (2008).
4. Choi, K., Kennedy, M., Kazarov, A., Papadimitriou, J. C. & Keller, G. A common precursor for hematopoietic and endothelial cells. *Development* **125,** 725–732 (1998).
5. de Bruijn, M. F. *et al.* Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta. *Immunity* **16,** 673–683 (2002).
6. Jaffredo, T., Gautier, R., Eichmann, A. & Dieterlen-Lievre, F. Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. *Development* **125,** 4575–4583 (1998).
7. Jaffredo, T. *et al.* From hemangioblast to hematopoietic stem cell: an endothelial connection? *Exp. Hematol.* **33,** 1029–1040 (2005).
8. Nishikawa, S. I. *et al. In vitro* generation of lymphohematopoietic cells from endothelial cells purified from murine embryos. *Immunity* **8,** 761–769 (1998).
9. Ueno, H. & Weissman, I. L. Clonal analysis of mouse development reveals a polyclonal origin for yolk sac blood islands. *Dev. Cell* **11,** 519–533 (2006).
10. Schroeder, T. Imaging stem cell driven mammalian regeneration. *Nature* **453,** 345–351 (2008).
11. Fraser, S. T. *et al.* Definitive hematopoietic commitment within the embryonic vascular endothelial-cadherin(+) population. *Exp. Hematol.* **30,** 1070–1078 (2002).
12. Sugiyama, D. *et al.* Erythropoiesis from acetyl LDL incorporating endothelial cells at the preliver stage. *Blood* **101,** 4733–4738 (2003).
13. Schroeder, T. *et al.* Recombination signal sequence-binding protein Jκ alters mesodermal cell fate decisions by suppressing cardiomyogenesis. *Proc. Natl Acad. Sci. USA* **100,** 4018–4023 (2003).
14. Schroeder, T. *et al.* Activated Notch1 alters differentiation of embryonic stem cells into mesodermal cell lineages at multiple stages of development. *Mech. Dev.* **123,** 570–579 (2006).
15. Hirashima, M., Kataoka, H., Nishikawa, S., Matsuyoshi, N. & Nishikawa, S. Maturation of embryonic stem cells into endothelial cells in an *in vitro* model of vasculogenesis. *Blood* **93,** 1253–1263 (1999).
16. Keller, G. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes Dev.* **19,** 1129–1155 (2005).
17. Nakano, T., Kodama, H. & Honjo, T. *In vitro* development of primitive and definitive erythrocytes from different precursors. *Science* **272,** 722–724 (1996).
18. Huber, T. L., Kouskoff, V., Fehling, H. J., Palis, J. & Keller, G. Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. *Nature* **432,** 625–630 (2004).
19. North, T. E. *et al.* Prostaglandin E2 regulates vertebrate haematopoietic stem cell homeostasis. *Nature* **447,** 1007–1011 (2007).
20. Kataoka, H. *et al.* Expressions of PDGF receptor alpha, c-Kit and Flk1 genes clustering in mouse chromosome 5 define distinct subsets of nascent mesodermal cells. *Dev. Growth Differ.* **39,** 729–740 (1997).
21. Shalaby, F. *et al.* A requirement for Flk1 in primitive and definitive hematopoiesis and vasculogenesis. *Cell* **89,** 981–990 (1997).
22. Yamashita, J. *et al.* Flk1-positive cells derived from embryonic stem cells serve as vascular progenitors. *Nature* **408,** 92–96 (2000).
23. Guo, R., Sakamoto, H., Sugiura, S. & Ogawa, M. Endothelial cell motility is compatible with junctional integrity. *J. Cell. Physiol.* **211,** 327–335 (2007).
24. Lampugnani, M. G. *et al.* A novel endothelial-specific membrane protein is a marker of cell–cell contacts. *J. Cell Biol.* **118,** 1511–1522 (1992).
25. Voyta, J. C., Via, D. P., Butterfield, C. E. & Zetter, B. R. Identification and isolation of endothelial cells based on their increased uptake of acetylated-low density lipoprotein. *J. Cell Biol.* **99,** 2034–2040 (1984).
26. Morita, K., Sasaki, H., Furuse, M. & Tsukita, S. Endothelial claudin: claudin-5/TMVCF constitutes tight junction strands in endothelial cells. *J. Cell Biol.* **147,** 185–194 (1999).
27. Taoudi, S. *et al.* Extensive hematopoietic stem cell generation in the AGM region via maturation of VE-cadherin^{+}CD45^{+} pre-definitive HSCs. *Cell Stem Cell* **3,** 99–108 (2008).
28. Ferkowicz, M. J. *et al.* CD41 expression defines the onset of primitive and definitive hematopoiesis in the murine embryo. *Development* **130,** 4393–4403 (2003).
29. Mikkola, H. K., Fujiwara, Y., Schlaeger, T. M., Traver, D. & Orkin, S. H. Expression of CD41 marks the initiation of definitive hematopoiesis in the mouse embryo. *Blood* **101,** 508–516 (2003).
30. Mitjavila-Garcia, M. T. *et al.* Expression of CD41 on hematopoietic progenitors derived from embryonic hematopoietic cells. *Development* **129,** 2003–2013 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to S. Nishikawa, M. A. Rieger, A. Hermann and M. Yoder for technical advice and discussions, M. Ogawa and H. Niwa for the VEVC2 and EB3 ESC lines, respectively, B. Schauberger for programming contributions, and A. Roth and C. Raithel for technical support. We thank M. Goetz, A. Hermann, M. A. Rieger and A. IJpenberg for critical reading of the manuscript. Part of this study was financed by the Deutsche Forschungsgemeinschaft to T.S. and by the Leading Project for the Realization of Regenerative Medicine to S.-I.N.

Author Contributions H.M.E. planned and performed experiments; S.-I.N. discussed results and commented on the manuscript; T.S. designed the study and experiments, developed the time-lapse imaging and cell-tracking technology, performed initial experiments, analysed data with H.M.E. and wrote the paper with H.M.E.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.S. (timm.schroeder@helmholtz-muenchen.de).

METHODS

Alexa Fluor 647 and Dylight488 labelling of antibodies was performed according to manufacturers' instructions (Bioprobes/Invitrogen, catalogue number A20186, and Thermo Scientific, product number 46402). Smooth muscle actin staining was performed as described earlier[14].

At day 4 of ESC differentiation, 1,440 living Flk-1^{+}E-cadherin^{-} cells per cm^2 were seeded on OP9 cells. The culture medium[13] was supplemented with 100 ng ml^{-1} stem cell factor (catalogue number 250-03, PeproTech) and 50 ng ml^{-1} 1,1′-dioctadecyl-3,3,3′,3′-tetramethyl-indocarbocyanine perchlorate (DiI)–Ac-LDL (BT-902, Biomedical Technologies).

Staging and dissection of CD-1 wild-type embryos was performed as described previously[11]. Embryos were collected in PBS, centrifuged at 300*g* for 5 min and resuspended in 1 ml 0.05% preheated trypsin/EDTA (GIBCO, Invitrogen). After incubation for 7 min at 37 °C the reaction was stopped by adding 9 ml Stop-solution containing 0.05 mg ml^{-1} DNase I (Sigma), 0.1% 1 M $CaCl_2 \times 2H_2O$ (Sigma), 20% FCS in HBSS (GIBCO, Invitrogen). After sedimentation, cells were incubated in ESC differentiation medium for 1 h at 37 °C and 5% CO_2. Cells were washed and stained with anti-VE-cadherin, anti-Flk-1 and anti-CD41 antibodies as well as propidium iodide for FACS sorting. Two-thousand sorted cells were seeded per cm^2 of confluent OP9 in ESC differentiation medium containing 100 U ml^{-1} SCF and incubated for 1 h at 37 °C and 5% CO_2 before starting the time-lapse experiment.

Phase-contrast pictures were acquired every 4 to 5 min using ×5 or ×10 phase-contrast objectives, and an AxioCamHRm camera (at 1,388 × 1,040 or 2,776 × 2,080 pixel resolutions) using Zeiss AxioVision 4.5 software. Mercury (HBO 103 W/2) or Xenon (XBO 75 W/2 OFR, both Osram) lamps were used for fluorescence illumination. Venus, DiI-labelled Ac-LDL and APC-labelled antibodies were detected using Zeiss filter sets 46HE, 43HE and 50, respectively. Movies for presentation were assembled using QuickTime software. Single-cell tracking was performed using self-written software on Siemens Celsius workstations and SUSE Linux operating systems. The software allows loading of pictures of a time-lapse acquisition with any temporal interval for all wavelength channels separately. All individual frames of time-lapse acquisitions (90% quality JPG for phase contrast, 100% JPG for fluorescence pictures) were displayed without further compression and separately for each wavelength to maintain image quality. Image contrast was manually enhanced for optimal recognition of relevant cellular features separately for each wavelength channel. Individual cells were observed and tracked manually. All relevant properties and behaviour (division, death, position, adherence, fluorescence, morphology, and so on) of cells of interest were stored and displayed in pedigrees as shown in Figs 1–4. All cell tracking was done by scientists; the presented analysis does not rely on data generated by an unsupervised computer algorithm for automated tracking. Only cells with unequivocal identity that could clearly be identified were used for analysis, and all cells with questionable identity were excluded from relevant analyses.

LETTERS

Calcium flickers steer cell migration

Chaoliang Wei[1], Xianhua Wang[1], Min Chen[1], Kunfu Ouyang[1], Long-Sheng Song[2] & Heping Cheng[1]

Directional movement is a property common to all cell types during development and is critical to tissue remodelling and regeneration after damage[1–3]. In migrating cells, calcium has a multifunctional role in directional sensing, cytoskeleton redistribution, traction force generation, and relocation of focal adhesions[1,4–7]. Here we visualize high-calcium microdomains ('calcium flickers') and their patterned activation in migrating human embryonic lung fibroblasts. Calcium flicker activity is dually coupled to membrane tension (by means of TRPM7, a stretch-activated Ca^{2+}-permeant channel of the transient receptor potential superfamily[8]) and chemoattractant signal transduction (by means of type 2 inositol-1,4,5-trisphosphate receptors). Interestingly, calcium flickers are most active at the leading lamella of migrating cells, displaying a 4:1 front-to-rear polarization opposite to the global calcium gradient[6]. When exposed to a platelet-derived growth factor gradient perpendicular to cell movement, asymmetric calcium flicker activity develops across the lamella and promotes the turning of migrating fibroblasts. These findings show how the exquisite spatiotemporal organization of calcium microdomains can orchestrate complex cellular processes such as cell migration.

In addition to extracellular chemoattractant stimuli, directional cell movement depends on an intracellular calcium signal that is well-organized in space, time and concentration[6,7,9]. Over a decade ago, it was shown that intracellular calcium displays a rear-to-front gradient, with the lowest concentration at the front of a migrating cell[6]. However, this observation seems to be paradoxical, because the leading lamella—the signalling and motility centre of a migrating cell—contains numerous effector proteins that require high levels of calcium for activation[10–13]. Although transient increases of calcium concentration have recently been observed in migrating cells, they are infrequent and mainly localized to the tail of the cell, and are thought to facilitate intermittent rear retraction[7]. Biochemical studies indicate that calcium entry is required to maintain ruffling structure, actin polymerization and the phosphatidylinositol-3,4,5-trisphosphate ($PtdIns(3,4,5)P_3$) signalling at the leading edge of macrophages[14]. So far, it has been perplexing how calcium regulates lamella dynamics during cell migration.

Using human embryonic lung fibroblasts (WI-38) as a model, we characterized the spatiotemporal organization of intracellular calcium signals with the aid of real-time confocal microscopy. In migrating WI-38 fibroblasts that overtly displayed leading and trailing edges, we detected a shallow decreasing gradient of global calcium concentration (indexed by the Fluo-4 to Fura red fluorescence ratio) that ran from the rear to the front (Fig. 1c, d), in agreement with previous findings[6,9]. Surprisingly, we found that discrete, local and short-lived high-calcium microdomains or 'calcium flickers', analogous to calcium sparks and puffs[15], occurred against a quiescent background (Supplementary Video). High-resolution linescan imaging revealed that the flickers occurred at a nearly constant rate of 1.92 ± 0.21 Hz per 100-μm linescan ($n = 18$; Fig. 1c). Individual events rapidly rose to about double the Fluo-4 fluorescence ($\Delta F/F_o = 1.16 \pm 0.02$, $n = 1{,}071$), lasted variably from 10 ms to 4 s, and were confined to an area 5.27 ± 0.05 μm in diameter (Fig. 1c). Importantly, calcium flickers were abundant at the leading lamella (Fig. 1a), including in motile lamellipodia (Fig. 1b), but were sharply reduced elsewhere, resulting in an approximately 4:1 front-to-rear polarization (Fig. 1d) that was opposite to the aforementioned global calcium gradient. Polarization of flicker activity was common to fibroblasts undergoing migration, but was not seen in stationary

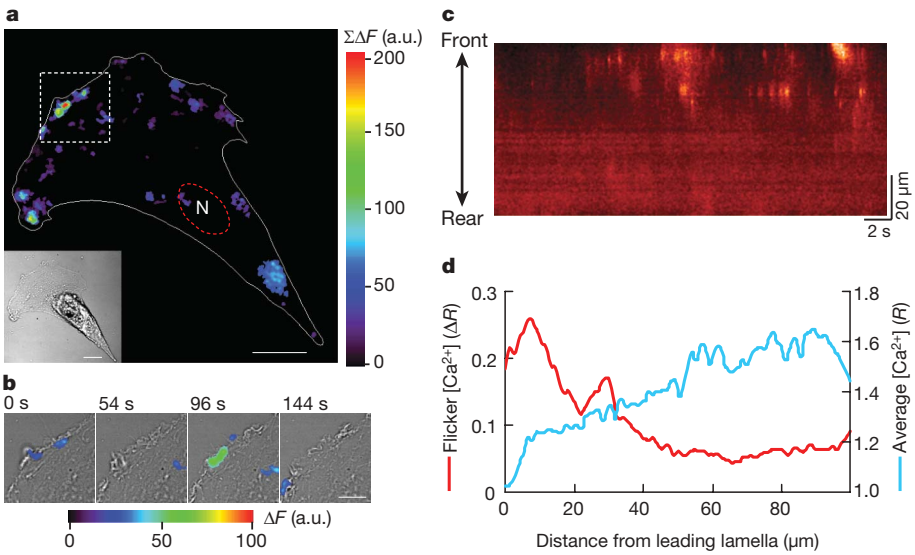


Figure 1 | Calcium flickers in migrating fibroblasts. **a**, Calcium flickers. In a polarized WI-38 fibroblast (inset), local calcium increases ($\Sigma\Delta F$) were summed over 30 consecutive images acquired at 6-s intervals. 'N' marks the nucleus. Scale bars, 15 μm. **b**, Calcium flickers (colour overlay) in motile lamellipodia (box in **a**). Scale bar, 5 μm; a.u., arbitrary units. **c**, Polarization of calcium flicker activity. The image consists of 10,000 front-to-rear linescans, expressed as the ratio between Fluo-4 and Fura red fluorescence (R). **d**, Opposing gradients of calcium flicker activity (ΔR) and global calcium (R, inclusive of flicker activity). Similar results were obtained in eight cells.

[1]Institute of Molecular Medicine, State Key Laboratory of Biomembrane and Membrane Biotechnology, Peking University, Beijing 100871, China. [2]Department of Internal Medicine, Division of Cardiovascular Medicine, University of Iowa Carver College of Medicine, Iowa City, Iowa 52242, USA.

fibroblasts lacking morphological polarity and displaying a lower flicker incidence (0.57 ± 0.10 Hz per 100-μm linescan, $n = 12$, $P < 0.05$ versus migrating cells; Supplementary Fig. 1). Thus, flickers represent a distinctive, heretofore unappreciated, modality of calcium signalling in migrating fibroblasts. Similar flickers were also evident in rat neonatal cardiac fibroblasts and 3T3-Swiss albino mouse embryonic fibroblasts (Supplementary Fig. 2).

In search of the molecular basis of calcium flickers, we showed that calcium influx through the stretch-activated cation channel (SACC) was obligatory. Application of Ca^{2+}-free medium containing 5 mM EGTA or streptomycin (200 μM), a SACC blocker[16], immediately abolished flicker activity in WI-38 fibroblasts (Fig. 2a). On average, the signal mass of calcium flickers (space–time integral of the flicker signal) decreased by $98.3 \pm 0.8\%$ (EGTA; $n = 9$, $P < 0.01$ versus control) or $93.1 \pm 1.3\%$ (streptomycin; $n = 6$, $P < 0.01$ versus control). Likewise, Gd^{3+} (200 μM), a non-specific SACC blocker[16], diminished the signal mass by $91.9 \pm 1.2\%$ ($n = 6$, $P < 0.01$ versus control). To determine whether mechanical forces can directly trigger flickers, we showed that shear stress applied to the front of migrating fibroblasts immediately evoked a flurry of flicker activity (Fig. 2e). In a different approach, relaxing or stretching the cell by pushing or pulling the flexible substrate (with a needle tip) suppressed or enhanced the flicker activity, respectively (Supplementary Fig. 3). Under cell-attached patch-clamp conditions[17], sudden application of negative pressure (~40 mm Hg) elicited bursting single-channel activity, whereas simultaneous confocal imaging visualized corresponding flicker-like events beneath the patch membrane (Fig. 2b). These results indicate that calcium flickers are triggered by calcium influx through SACCs.

SACCs belong to the transient receptor potential (TRP) ion channel superfamily. In mammals, about eight TRP channels in four subfamilies are thought to be sensitive to mechanical forces while being calcium-permeable[8]. Of these, there were relatively high *TRPM7*, *TRPC6*, *TRPV2* and *PKD2* (also known as PC2 or TRPP2) messenger RNA levels in WI-38 fibroblasts (Supplementary Fig. 4). Using RNA interference (RNAi), we found that calcium flickers were virtually abolished by ~75% knockdown of *TRPM7*, but not by that of the other three TRP channels (Fig. 2c, d and Supplementary Fig. 5). More importantly, similar shear stress was unable to evoke flickers in *TRPM7* knockdown cells, which displayed rare basal flicker activity (Fig. 2e). These data pinpointed TRPM7 as the specific SACC responsible for transducing mechanical signals into calcium flickers. A characteristic of TRPM7 is its sensitivity to inhibition by Mg^{2+} in addition to Gd^{3+} (ref. 18). Increasing extracellular Mg^{2+} from 1.0 mM to 10 mM largely abolished calcium flickers, whereas removing it enhanced flicker production (Supplementary Fig. 6). Thus, TRPM7 acts as the mechanical sensor, the calcium flicker igniter and the mechanochemical transducer in fibroblasts, revealing a previously unknown role of this SACC in the regulation of cell migration (see below).

Because calcium release from the endoplasmic reticulum amplifies calcium influxes by means of the calcium-induced calcium release mechanism[15], we next investigated whether endoplasmic reticulum calcium release participates in calcium flicker production. Inhibition of endoplasmic reticulum calcium recycling with the Ca^{2+} ATPase inhibitor thapsigargin (5 μM, 20 min incubation, after recession of the initial calcium transient) halved flicker amplitude without affecting flicker probability (Fig. 3a, c). Similar results were obtained by inhibiting the inositol-1,4,5-trisphosphate ($Ins(1,4,5)P_3$) receptor ($Ins(1,4,5)P_3R$) with xestospongin C (5 μM), whereas $Ins(1,4,5)P_3$-BM (2 μM), a membrane-permeable ester precursor of $Ins(1,4,5)P_3$ (ref. 19), enhanced the flicker amplitude, and ryanodine receptor inhibition (ryanodine, 25 μM) had no significant effect (Fig. 3a, c). Quantitative real-time polymerase chain reaction (PCR) results showed that type 2 $Ins(1,4,5)P_3R$ ($Ins(1,4,5)P_3R2$) and type 3 $Ins(1,4,5)P_3R$ ($Ins(1,4,5)P_3R3$) are the primary $Ins(1,4,5)P_3R$ isoforms expressed in WI-38 fibroblasts (Supplementary Fig. 7). In contrast to TRPM7, RNA interference knockdown of the gene encoding $Ins(1,4,5)P_3R2$ (~80%), but not that encoding $Ins(1,4,5)P_3R3$ (~60%), significantly decreased flicker amplitude but failed to alter flicker probability (Fig. 3b, c). This $Ins(1,4,5)P_3R$ isoform specificity is consistent with the fact that $Ins(1,4,5)P_3R2$ is more sensitive to $Ins(1,4,5)P_3$ and displays little calcium-dependent inactivation[20]. Taken these findings together, we concluded that calcium entry by means of TRPM7 is locally amplified by calcium release through $Ins(1,4,5)P_3R2$ in the event of a calcium flicker. Coupling $Ins(1,4,5)P_3R2$ to TRPM7 would enable flicker activity to decode $Ins(1,4,5)P_3$-linked chemoattractant signal transduction.

Given the role of TRPM7 as a mechanical sensor and a calcium flicker igniter, we anticipated that flicker activity would be coupled to

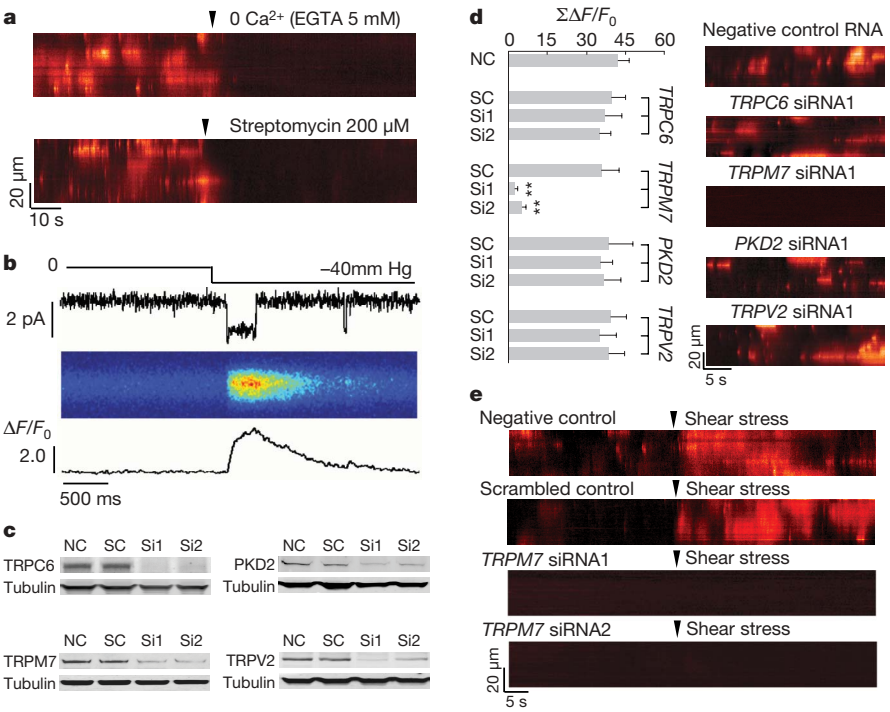


Figure 2 | Triggering calcium flickers by TRPM7. **a**, Abolition of calcium flickers by streptomycin or removal of external calcium. **b**, Visualization of calcium entry through single SACCs. From top to bottom: suction through the patch pipette; single-channel currents; and linescan image and line plot of local calcium transients. **c**, RNAi knockdown of *TRPC6*, *TRPM7*, *TRPV2* and *PKD2* assayed by western blotting. NC, negative control RNA; SC, scrambled control RNA; Si1 and Si2, different siRNA sequences (see Supplementary Table 2). **d**, Knockdown of *TRPM7*, but not *TRPC6*, *TRPV2* or *PKD2*, abolished calcium flickers. Data are expressed as mean and s.e.m.; $n = 12$–21 cells in each group; double asterisk, $P < 0.01$ versus the respective SC. **e**, *TRPM7* knockdown prevented shear-stress-induced calcium flickers.

the migration-associated traction force. Indeed, the map of flicker ignition sites at the front of a cell largely overlapped, although with subtle differences, the matrix of focal adhesions (Fig. 4a and Supplementary Fig. 8), where traction force is created and transmitted[21]. Rapid local application of RGDS (2 mM, 1 min), which contains the RGD sequence that is recognized by integrins[22], enhanced the flicker activity, whereas the control peptide RGES was ineffective (Fig. 4b, c), consistent with the finding that RGDS stimulates calcium transients in neuronal filopodia and growth cones[23]. Disruption of protrusion by transient frontal application of cytochalasin D and inhibition of myosin ATPase by 2,3-butanedione monoxime or (−)blebbistatin[24] both inhibited lamella flicker production (Fig. 4b, c). These lines of evidence support the idea that decoding local membrane tension by flicker activity depends on cytoskeletal and morphological integrity.

Despite the low global calcium concentration, high-calcium microdomains created by mechanical stress in the leading lamella

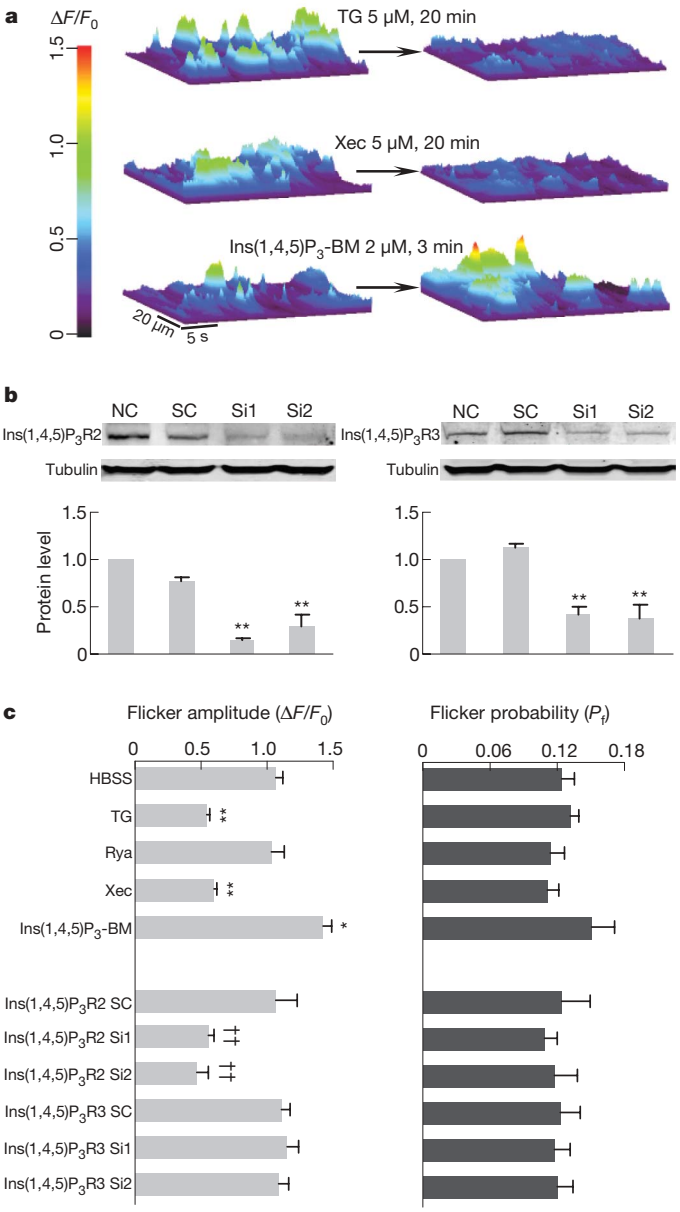


Figure 3 | Amplifying TRPM7 calcium flickers by store calcium release through type 2 Ins(1,4,5)P_3 receptors. **a**, Typical calcium flicker responses to thapsigargin (TG), xestospongin C (Xec) or Ins(1,4,5)P_3-BM. Linescan data of $\Delta F/F_0$ are rendered as surface plots. **b**, Western blotting of Ins(1,4,5)P_3Rs. See Supplementary Table 2 for sequences of negative control (NC), respective scrambled control (SC), Si1 and Si2. Data are expressed as mean and s.e.m. ($n = 3$–4). Double asterisk, $P < 0.01$ versus the respective SC. **c**, Flicker amplitude and probability under different experimental conditions. P_f, fractional time spent in calcium flickers; HBSS, HEPES-buffered saline solution; Rya, ryanodine (25 μM, 15 min). $n = 10$–29; asterisk, $P < 0.05$; double asterisk, $P < 0.01$ versus HBSS; double dagger, $P < 0.01$ versus the respective SC. Note that perturbing the endoplasmic reticulum calcium release mainly altered flicker amplitude without affecting flicker probability.

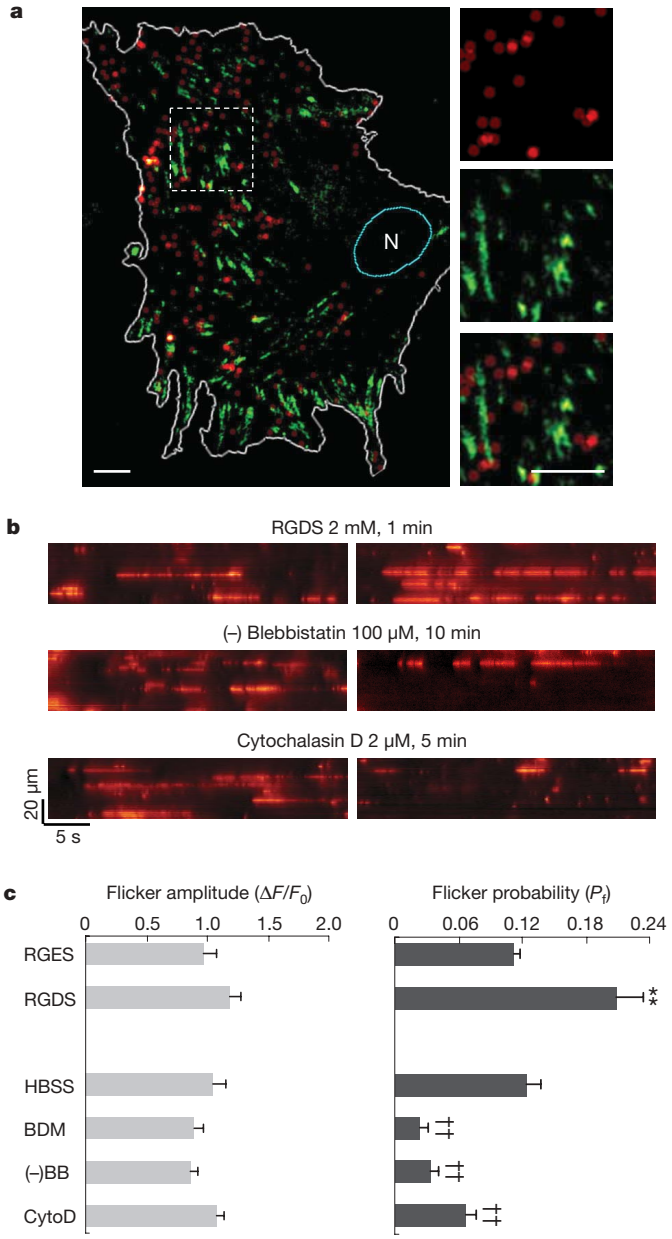


Figure 4 | Traction force generation and calcium flicker activity. **a**, Maps for calcium flicker ignition sites (red dots) and focal adhesions (green). Focal adhesions were visualized by immunostaining for integrin α5 after calcium flicker acquisition. Enlarged views of calcium flickers, focal adhesions and their overlay are shown to the right (top, middle and bottom, respectively). 'N' denotes the nucleus. Scale bar, 8 μm. **b**, Lamella flicker before (left) and after (right) application of compounds that affect traction-force-generating elements. **c**, Statistics of calcium flicker amplitude and probability (P_f). (−)BB, (−)blebbistatin (100 μM); BDM, 2,3-butanedione monoxime (10 mM); cytoD, cytochalasin D; HBSS, HEPES-buffered saline solution. Error bars represent s.e.m.; $n = 10$–19 in each group. Double asterisk, $P < 0.01$ versus RGES; double dagger, $P < 0.01$ versus HBSS. Note that flicker probability rather than amplitude was preferentially altered by varying the traction force, in contrast to the situation shown in Fig. 3.

may activate a multitude of local calcium-dependent events critical to cell polarization and movement, including the PtdIns(3,4,5)P_3 signalling cascade[25,26], a parallel phospholipase A2-mediated signalling mechanism[27], cytoskeleton dynamics such as actin remodelling[10], focal adhesion detachment and relocation, and actin–myosin contraction. Next, we sought to determine the physiological role of calcium flickers in regulating cell migration, particularly turning behaviour that is almost entirely carried out within the leading lamella.

Platelet-derived growth factor (PDGF) is a well known chemoattractant that stimulates fibroblast migration during wound healing[3]. Its intracellular signalling pathways include generating traction force by Rac-dependent protrusion[28] and activation of the phospholipase C–PtdIns(4,5)P_2–Ins(1,4,5)P_3 signalling cascade[29], both convergent on flicker production. When migrating fibroblasts were exposed to uniform PDGF (0.8 nM), increases in both flicker amplitude and probability were accompanied by a decrease in directional persistence (Supplementary Fig. 9), the latter suggesting an enhanced propensity for turning of the cell[30]. When a PDGF gradient was applied in the direction perpendicular to cell movement, migrating fibroblasts no longer moved persistently along the original path; instead, they turned towards the higher PDGF concentration (Fig. 5a). Time-lapse imaging revealed a rapid increase in lamella flicker activity and an accentuated front-to-rear polarization (Fig. 5a). More importantly, a greater flicker signal mass (SM) was found in the portion facing the PDGF source (SM_α) than in the portion further away from the source (SM_β), indicating the development of an asymmetry of flicker activity within the lamella (Fig. 5a). To demonstrate a linkage between lamella flicker asymmetry and turning behaviour, we examined the cumulative difference between SM_α and SM_β ($\Sigma SM_{\alpha-\beta}$) and found that the time course of $\Sigma SM_{\alpha-\beta}$ nearly overlapped with that for the distance travelled in the direction of the PDGF gradient (*y*-distance; Fig. 5b). On average, the correlation coefficient between $\Sigma SM_{\alpha-\beta}$ and *y*-distance was 0.72 in 25 migrating fibroblasts. Hence, patterned flicker activation in the leading lamella may translate directional sense and steer the cell to turn in response to PDGF gradients.

To test this hypothesis further, we showed that impaired PDGF-induced lamella flicker asymmetry or reduction of $\Sigma SM_{\alpha-\beta}$ by streptomycin and *TRPM7* or Ins(1,4,5)P_3R2 knockdown was accompanied by dwindling of the turning angle in a roughly proportional manner (Fig. 5c). Likewise, a robust match between the flicker activity and population chemotaxis was revealed by various pharmacological and molecular interventions including *TRPM7* and Ins(1,4,5)P_3R2 knockdown, and SACC blockade (Fig. 5d and Supplementary Fig. 10). Furthermore, loading EGTA ester to disturb the flicker signal (Supplementary Fig. 11) compromised the turning and chemotaxis abilities (Fig. 5c, d), whereas flicker activation by Ins(1,4,5)P_3-BM enhanced both of these (Fig. 5c, d), suggesting a causal link between calcium flicker activity and fibroblast turning and chemotaxis.

We have demonstrated that calcium flickers arising from TRPM7 and Ins(1,4,5)P_3R2 have an essential role in steering migrating

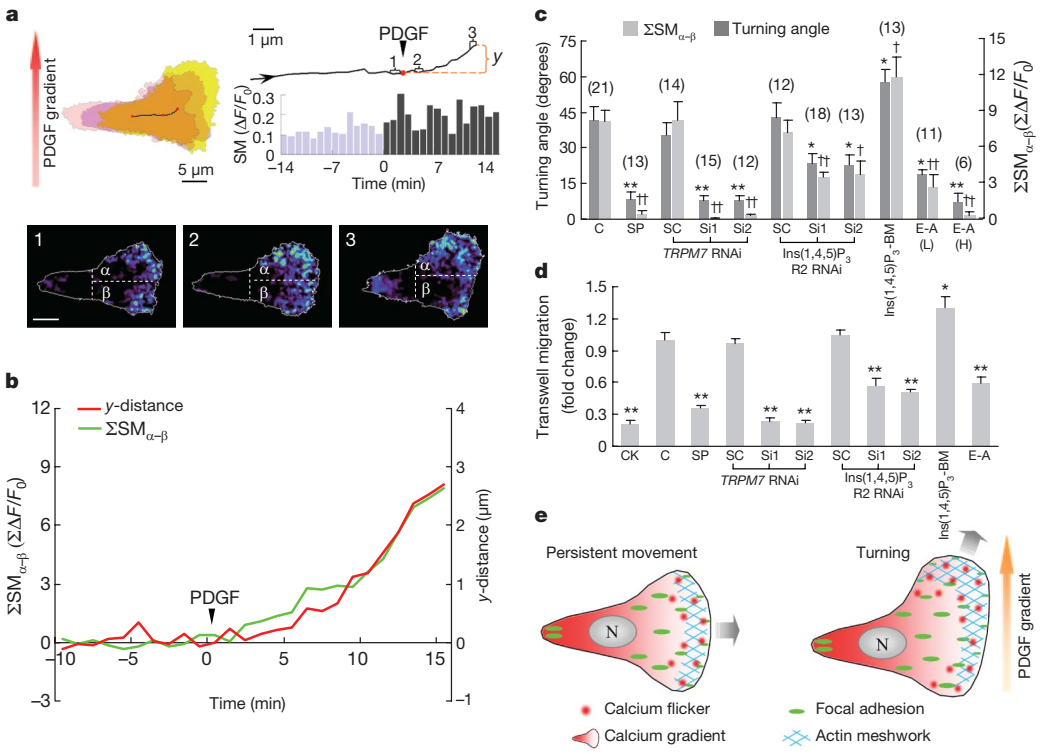


Figure 5 | Calcium flickers steer fibroblast turning. **a**, Asymmetrical lamella flicker activity induced by a PDGF gradient. Top left, contours of the cell border at −14 min (pink), 0 min (purple) and 15 min (yellow). Top right, trajectory of the centre of the cell and time course of calcium flicker production (bar graph). SM, signal mass of calcium flickers within the lamella. Bottom panels: overlays of calcium flickers in 1-min windows (labelled 1–3 in the trajectory above). Dashed lines bisect the leading lamella into upper (α, facing the PDGF source) and lower portions (β). **b**, Correlation between cumulative asymmetric flicker activity ($\Sigma SM_{\alpha-\beta}$) and displacement along the PDGF gradient (*y*-distance). **c**, Relationship between turning angle and $\Sigma SM_{\alpha-\beta}$ (at 15 min). C, control with no treatment; SP, streptomycin, 200 μM; SC, scrambled control; Si1 and Si2, two siRNA duplexes; Ins(1,4,5)P_3-BM, 2 μM; E-A (L), EGTA-AM, 2 μM; E-A (H), EGTA-AM, 20 μM. Data are expressed as mean and s.e.m.; *n* values are shown in parentheses. Asterisk, $P < 0.05$; double asterisk, $P < 0.01$ versus turning angle of control or respective SC; dagger, $P < 0.05$; double dagger, $P < 0.01$ versus $\Sigma SM_{\alpha-\beta}$ of control or respective SC. **d**, Chemotaxis of WI-38 fibroblasts. CK, chemokinesis assay with the same concentration of PDGF-BB on both sides of the well; E-A, EGTA-AM (2 μM). $n = 4$–10; asterisk, $P < 0.05$; double asterisk, $P < 0.01$ versus control or respective SC. **e**, Cartoons of patterned calcium flicker activation in persistently moving (left) and turning (right) fibroblasts. Calcium flicker activity during persistent movement displays a front-to-rear polarization, opposing a rear-to-front global calcium gradient. During turning, calcium flicker activity becomes asymmetric across the leading lamella, in addition to enhanced flicker frequency and accentuated front-to-rear polarization.

fibroblasts. Despite the observation that the global calcium gradient is opposite to the direction of cell migration, high calcium flicker activity would enable activation of calcium signalling cascades amidst a low-calcium background at the leading edge, such that spatiotemporally patterned calcium flicker activity can orchestrate the complex turning behaviour of migrating cells (Fig. 5e). The coupling of TRPM7-mediated force-transducing calcium influx and local $Ins(1,4,5)P_3$-induced calcium release would make this an ideal system for locomotion in response to chemoattractants (Fig. 5e). The present finding may have general ramifications because growth cones of neurons turn away from the side at which the filopodia displays higher local calcium signals[23] and calcium influx is essential for maintaining the leading-edge structure and activity in macrophages[14]. As such, unveiling calcium flickers in migrating cells opens a new avenue to investigate how local calcium signals orchestrate diverse biochemical pathways in the guidance of directional movement.

METHODS SUMMARY

Human lung embryonic WI-38 fibroblasts were obtained from the American Type Culture Collection and were loaded with Fluo-4 AM alone or in combination with Fura red AM. Images were captured on a Zeiss LSM 510 confocal microscope. The primers used for PCR with reverse transcription (RT–PCR) or quantitative RT–PCR are shown in Supplementary Table 1. Short interfering RNA (siRNA) sequences and scrambled controls were designed using RNAi Designer (http://www.invitrogen.com/rnai; Supplementary Table 2). WI-38 fibroblasts were transfected with 100 nM siRNA duplexes using Lipofectamine RNAiMax (Invitrogen) according to the manufacturer's recommendations. Western blotting or functional studies were carried out 72 h after transfection. Fibroblasts with an overt leading lamella and a thin trailing edge were selected for the cell-turning study. A PDGF-BB-containing micropipette was placed ~150 μm away from one side of the cell to establish a PDGF-BB gradient. Twenty-four-well Transwell plates (Corning) were used for the chemotaxis assay. The outer wells contained 0.8 nM PDGF-BB and overnight-starved WI-38 fibroblasts in PDGF-BB-free medium. Designated drugs were added to the inserts. For the assay, cells were loaded with calcein AM and then fixed immediately with formaldehyde. Cells under the lower surface of the polycarbonate membrane were imaged and analysed. For electrophysiological measurements, the cell-attached patch-clamp technique, using an EPC-7 amplifier, was applied to fibroblasts preloaded with Fluo-4 AM. To activate SACCs, mechanical suction of ~40 mm Hg was applied by means of a syringe connected to the patch pipette. In the mechanical perturbation experiment, shear stress was locally applied by a gentle jet flow by means of a patch pipette ~80 μm from the front of migrating fibroblasts. Digital image processing used IDL software (Research Systems) and custom-devised computer algorithms. Statistical data are expressed as mean and s.e.m. Student's *t*-test and paired *t*-test were applied when appropriate.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 13 April; accepted 21 October 2008.
Published online 31 December 2008.

1. Ridley, A. J. *et al.* Cell migration: integrating signals from front to back. *Science* **302,** 1704–1709 (2003).
2. Martin, P. & Parkhurst, S. M. Parallels between tissue repair and embryo morphogenesis. *Development* **131,** 3021–3034 (2004).
3. Werner, S. & Grose, R. Regulation of wound healing by growth factors and cytokines. *Physiol. Rev.* **83,** 835–870 (2003).
4. Van Haastert, P. J. & Devreotes, P. N. Chemotaxis: signalling the way forward. *Nature Rev. Mol. Cell Biol.* **5,** 626–634 (2004).
5. Pettit, E. J. & Fay, F. S. Cytosolic free calcium and the cytoskeleton in the control of leukocyte chemotaxis. *Physiol. Rev.* **78,** 949–967 (1998).
6. Brundage, R. A., Fogarty, K. E., Tuft, R. A. & Fay, F. S. Calcium gradients underlying polarization and chemotaxis of eosinophils. *Science* **254,** 703–706 (1991).
7. Lee, J., Ishihara, A., Oxford, G., Johnson, B. & Jacobson, K. Regulation of cell movement is mediated by stretch-activated calcium channels. *Nature* **400,** 382–386 (1999).
8. Nilius, B., Owsianik, G., Voets, T. & Peters, J. A. Transient receptor potential cation channels in disease. *Physiol. Rev.* **87,** 165–217 (2007).
9. Hahn, K., DeBiasio, R. & Taylor, D. L. Patterns of elevated free calcium and calmodulin activation in living cells. *Nature* **359,** 736–738 (1992).
10. Stossel, T. P., Fenteany, G. & Hartwig, J. H. Cell surface actin remodeling. *J. Cell Sci.* **119,** 3261–3264 (2006).
11. Robinson, R. C. *et al.* Domain movement in gelsolin: a calcium-activated switch. *Science* **286,** 1939–1942 (1999).
12. Chew, T. L., Wolf, W. A., Gallagher, P. J., Matsumura, F. & Chisholm, R. L. A fluorescent resonant energy transfer-based biosensor reveals transient and regional myosin light chain kinase activation in lamella and cleavage furrows. *J. Cell Biol.* **156,** 543–553 (2002).
13. Franco, S. J. *et al.* Calpain-mediated proteolysis of talin regulates adhesion dynamics. *Nature Cell Biol.* **6,** 977–983 (2004).
14. Evans, J. H. & Falke, J. J. Ca^{2+} influx is an essential component of the positive-feedback loop that maintains leading-edge structure and activity in macrophages. *Proc. Natl Acad. Sci. USA* **104,** 16176–16181 (2007).
15. Cheng, H. & Lederer, W. J. Calcium sparks. *Physiol. Rev.* **88,** 1491–1545 (2008).
16. Hamill, O. P. & McBride, D. W. Jr. The pharmacology of mechanogated membrane ion channels. *Pharmacol. Rev.* **48,** 231–252 (1996).
17. Zou, H., Lifshitz, L. M., Tuft, R. A., Fogarty, K. E. & Singer, J. J. Visualization of Ca^{2+} entry through single stretch-activated cation channels. *Proc. Natl Acad. Sci. USA* **99,** 6404–6409 (2002).
18. Numata, T., Shimizu, T. & Okada, Y. TRPM7 is a stretch- and swelling-activated cation channel involved in volume regulation in human epithelial cells. *Am. J. Physiol. Cell Physiol.* **292,** C460–C467 (2007).
19. Thomas, D. *et al.* Microscopic properties of elementary Ca^{2+} release sites in non-excitable cells. *Curr. Biol.* **10,** 8–15 (2000).
20. Patterson, R. L., Boehning, D. & Snyder, S. H. Inositol 1,4,5-trisphosphate receptors as signal integrators. *Annu. Rev. Biochem.* **73,** 437–465 (2004).
21. Horwitz, A. R. & Parsons, J. T. Cell migration—movin' on. *Science* **286,** 1102–1103 (1999).
22. Pierschbacher, M. D. & Ruoslahti, E. Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. *Nature* **309,** 30–33 (1984).
23. Gomez, T. M., Robles, E., Poo, M. & Spitzer, N. C. Filopodial calcium transients promote substrate-dependent growth cone turning. *Science* **291,** 1983–1987 (2001).
24. Shu, S., Liu, X. & Korn, E. D. Blebbistatin and blebbistatin-inactivated myosin II inhibit myosin II-independent processes in Dictyostelium. *Proc. Natl Acad. Sci. USA* **102,** 1472–1477 (2005).
25. Price, L. S. *et al.* Calcium signaling regulates translocation and activation of Rac. *J. Biol. Chem.* **278,** 39413–39421 (2003).
26. Weiner, O. D. *et al.* A $PtdInsP_3$- and Rho GTPase-mediated positive feedback loop regulates neutrophil polarity. *Nature Cell Biol.* **4,** 509–513 (2002).
27. van Haastert, P. J., Keizer-Gunnink, I. & Kortholt, A. Essential role of PI3-kinase and phospholipase A2 in *Dictyostelium discoideum* chemotaxis. *J. Cell Biol.* **177,** 809–816 (2007).
28. Hawkins, P. T. *et al.* PDGF stimulates an increase in GTP–Rac via activation of phosphoinositide 3-kinase. *Curr. Biol.* **5,** 393–403 (1995).
29. Claesson-Welsh, L. Platelet-derived growth factor receptor signals. *J. Biol. Chem.* **269,** 32023–32026 (1994).
30. Ware, M. F., Wells, A. & Lauffenburger, D. A. Epidermal growth factor alters fibroblast migration speed and directional persistence reciprocally and in a matrix-dependent manner. *J. Cell Sci.* **111,** 2423–2432 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank I. C. Bruce, X. Zhu, J. J. Ma, H. Q. Cao, J. Liu, M. Zheng, X. D. Fu and R. P. Xiao for discussions, and M. Gorospe, Z. C. Liang, Q. Du, C. M. Cao, H. Huang, X. M. Lan, N. Lin, Y. R. Wang, R. S. Song and Y. Zhang for technical support. Special thanks to X. D. Fu for making his laboratory facility available to us. This work was supported by the Major State Basic Research Development Program of China (2007CB512100), and the National Natural Science Foundation of China (30630021 and 30800371) and an NIH grant (HL090905).

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to H.C. (chengp@pku.edu.cn) and C.W. (chaoliang.wei@gmail.com).

METHODS

Cell culture. Human lung embryonic WI-38 fibroblasts (21 population doublings) obtained from the American Type Culture Collection were maintained and subcultured to 28 population doublings in MEM (Gibco) supplemented with 10% FBS (Hyclone), 2 mM glutamine and 200 units ml^{-1} penicillin in a incubator with parameters preset at 37 °C and 5% CO_2. For functional experiments, cells were plated at a density of 1×10^4 per cm^2 and cultured for 10 h on coverslips coated with 5 μg ml^{-1} fibronectin (Sigma).

Calcium imaging. WI-38 cells were loaded with Fluo-4 AM (5 μM) alone or in combination with Fura red AM (5 μM) for 6 min at 37 °C, rinsed twice, and then bathed in HEPES-buffered saline solution containing (in mM): 134 NaCl, 5.4 KCl, 1.0 $MgSO_4$, 1.0 NaH_2PO_4, 1.8 $CaCl_2$, 20 HEPES and 5 D-glucose (pH 7.4) with 1% FBS, unless otherwise specified. Cells were placed in a 37 °C heated chamber (Zeiss S-Type incubator) and imaged on a Zeiss LSM 510 confocal microscope with a ×40 oil objective (NA 1.3) at radial and axial resolutions of 0.4 μm and 1.0 μm, respectively. For ratiometric imaging, cells were excited at 488 nm, emission was detected at 505–550 nm (Fluo-4 signal) and >633 nm (Fura red signal), and the differential interference contrast transmission image was acquired simultaneously. For migration path analysis and calcium flicker signal mass measurements, 300–600 time-lapse images were acquired at 6-s intervals. High-resolution linescan imaging of calcium flickers was performed at 3 ms per linescan.

PCR. Total RNA was isolated from WI-38 fibroblasts (28 population doublings) with TRI Reagent (Sigma) and converted to complementary DNA using M-MLV reverse transcriptase (Promega). Quantitative RT–PCR reactions were carried out using these cDNAs in an iQ5 real-time PCR detection system (BioRad). Results were read out using iQ5 optical system software. All samples showing primer dimer formation or spurious, non-specific peaks, as indicated by the dissociation curve, were excluded from analysis. The primers are shown in Supplementary Table 1.

RNA interference. RNAi sequences for Ins(1,4,5)P_3R isoforms and TRP channels were designed using RNAi Designer (http://www.invitrogen.com/rnai; Supplementary Table 2). Each scrambled control was designed corresponding to first duplex of siRNA. In brief, corresponding siRNA duplexes were synthesized (GenePharma or Invitrogen) and transfected into cells with Lipofectamine RNAiMax (Invitrogen) according to the manufacturer's recommendations. Western blotting or functional studies were carried out 72 h after transfection.

Western blotting. Total protein extracted from WI-38 cells with siRNA treatment was separated on 4–12% NuPAGE Novex Bis-Tris gels (Invitrogen) and transferred to PVDF membranes (Millipore). After blocking for 1 h with 5% non-fat dry milk, the PVDF membrane was probed with primary antibody (anti-Ins(1,4,5)P_3R2 was the gift of J. Chen; anti-Ins(1,4,5)P_3R3 was from Santa Cruz; anti-tubulin from Sigma; anti-TRPC6 from Millipore; anti-TRPV2 from ABR; and anti-PKD2 and anti-TRPM7 from Abcam) for 2 h at room temperature (20–25 °C), and then secondary antibody (IRDye-conjugated anti-mouse, anti-rabbit and anti-goat IgG from LI-COR) for 1 h at room temperature. Immunoblots were detected using the Odyssey imaging system.

Cell migration analysis. Fibroblasts with an overt leading lamella and a thin trailing edge were selected for migration analysis. The outer boundary of the cell was extracted from the respective fluorescence image for calculation of its centre of gravity. The centres of consecutive images (6 s apart) defined the trajectory of cell movement. Migration speed was calculated as the average displacement per min during 30 min. Directional persistence (D/T ratio) was calculated as the ratio between the linear displacement and the total length of the trajectory during 30 min.

To establish a PDGF-BB (PeproTech) gradient perpendicular to the long axis of a polarized migrating fibroblast, a 5 μm internal diameter, PDGF-BB-containing (3 nM) micropipette was placed ~150 μm away from one side of the cell. By visualization of sulphurhodamine fluorescence under similar conditions, we estimated an average PDGF concentration of 1 nM and an edge-to-edge difference of 0.4 nM across the leading lamella (~40 μm).

Chemotaxis assay. Twenty-four-well Transwell plates with inserts containing 8-μm pores in a polycarbonate membrane (Corning) were used for the chemotaxis assays. In brief, the outer wells contained 600 μl MEM medium containing 1% FBS with PDGF-BB (0.8 nM) as a chemoattractant. Approximately 8×10^3 overnight-starved (1% FBS) WI-38 fibroblasts in 100 μl PDGF-BB-free MEM medium containing 1% FBS and designated drug were added to each insert. In the chemokinesis control group, PDGF-BB (0.8 nM) was also added to the insert to abolish the concentration gradient. Transwell plate was then incubated for 12 h in an incubator with parameters preset at 37 °C and 5% CO_2 before assay.

For the assay, the inserts were loaded with 5 μM calcein AM for 10 min and then fixed immediately with 3% formaldehyde for 10 min. Cells in inserts were cleared and those under the lower surface of the polycarbonate membrane were imaged and analysed.

Application of mechanical forces. Shear stress was locally applied by a gentle jet flow (4 cm H_2O pressure) by means of a patch pipette (10 μm internal diameter) ~80 μm away from the front of migrating fibroblasts. Note that jet flow used in local drug delivery (1 cm H_2O pressure, ~120 μm placement, pipette with 5 μm internal diameter) did not alter calcium flicker activity ($n = 4$).

Recording SACC currents and imaging local calcium influx. Cell-attached patch-clamp technique, using an EPC-7 amplifier, was applied to fibroblasts preloaded with the calcium indicator, Fluo-4 AM. The patch pipette (2–3 MΩ) solution contained (in mM): NaCl 140, KCl 5.4, $MgCl_2$ 1.0, HEPES 20 and $CaCl_2$ 1.8 (pH 7.4, adjusted with NaOH). To activate SACCs, mechanical suction of ~40 mm Hg was applied by means of a syringe connected to the patch pipette while the patch membrane was held 80 mV more negative than the resting membrane potential to enhance Ca^{2+} entry. The single-channel currents were filtered at 3 kHz and digitised at 5 kHz with pClamp 6.0 software. Linescan images of local calcium immediately beneath the patch membrane were acquired simultaneously at 3 ms resolution.

Data analysis. Digital image processing used IDL software (Research Systems) and custom-devised computer algorithms. Statistical data are expressed as mean and s.e.m. Student's t-test and paired t-test were applied when appropriate. A P value less than 0.05 was considered as statistically significant.

Vol 457|12 February 2009|doi:10.1038/nature07575

LETTERS

ABIN-1 is a ubiquitin sensor that restricts cell death and sustains embryonic development

Shigeru Oshima[1]*, Emre E. Turer[1]*, Joseph A. Callahan[1], Sophia Chai[1], Rommel Advincula[1], Julio Barrera[1], Nataliya Shifrin[1], Bettina Lee[1], Benjamin Yen[1], Tammy Woo[1], Barbara A. Malynn[1] & Averil Ma[1]

Proteins that directly regulate tumour necrosis factor receptor (TNFR) signalling have critical roles in regulating cellular activation and survival. ABIN-1 (A20 binding and inhibitor of NF-κB) is a novel protein that is thought to inhibit NF-κB signalling[1,2]. Here we show that mice deficient for ABIN-1 die during embryogenesis with fetal liver apoptosis, anaemia and hypoplasia. ABIN-1 deficient cells are hypersensitive to tumour necrosis factor (TNF)-induced programmed cell death, and TNF deficiency rescues ABIN-1 deficient embryos. ABIN-1 inhibits caspase 8 recruitment to FADD (Fas-associated death domain-containing protein) in TNF-induced signalling complexes, preventing caspase 8 cleavage and programmed cell death. Moreover, ABIN-1 directly binds polyubiquitin chains and this ubiquitin sensing activity is required for ABIN-1's anti-apoptotic activity. These studies provide insights into how ubiquitination and ubiquitin sensing proteins regulate cellular and organismal survival.

To investigate the potential roles of ABIN-1 in regulating TNF signals, we targeted the gene that encodes ABIN-1, *Tnip1*, in embryonic stem cells to generate mice deficient for ABIN-1 (Supplementary Fig. 1a–c). Deletion of exons 12–15 of *Tnip1* eliminated the full length (approximately 80 kDa) ABIN-1 protein (Supplementary Fig. 1d). While $Tnip1^{+/-}$ mice appear normal, very few live born $Tnip1^{-/-}$ mice were obtained from intercrossed $Tnip1^{+/-}$ parents (Table 1). $Tnip1^{-/-}$ embryos were found in Mendelian ratios at post coital day (E)12.5 to E18.5, but were only rarely obtained as live born mice (Table 1).

To better understand how ABIN-1 sustains embryonic development, we noted that $Tnip1^{-/-}$ embryos were smaller and paler than $Tnip1^{+/+}$ littermates at E16.5 and E18.5 (Fig. 1a). Quantitation of fetal weights and haematocrits of these embryos revealed that $Tnip1^{-/-}$ embryos were significantly smaller and anaemic compared to control embryos (Fig. 1b, c). Histological analyses of multiple tissues at various stages of development revealed that fetal livers of E16.5 $Tnip1^{-/-}$ embryos contained multiple areas of hypocellularity (Fig. 1d). Immunohistochemical analyses of caspase 3 cleavage indicated that these areas contained large numbers of apoptotic cells that were not detected in control embryos (Fig. 1d). Thus, ABIN-1 is critical for preventing fetal liver apoptosis and sustaining fetal haematopoiesis.

Table 1 | ABIN-1 is required for regulating TNF signals *in utero*

TNF genotype	Age	$Tnip1^{+/+}$	$Tnip1^{+/-}$	$Tnip1^{-/-}$	$Tnip1^{-/-}$ (%)	Total no.
$Tnf^{+/+}$	E15.5	12	32	17	(27.9%)	61
$Tnf^{+/+}$	E16.5	6	29	11	(23.9%)	46
$Tnf^{+/+}$	E18.5	11	16	11	(28.9%)	38
$Tnf^{+/+}$	Newborn	12	28	1	(2.4%)	41
$Tnf^{+/+}$	3–4 week	37	92	4	(3.0%)	133
$Tnf^{+/-}$	Live born	19	28	4	(7.8%)	51
$Tnf^{-/-}$	Live born	27	61	22	(20%)	110

Numbers of embryos or live born pups of the indicated genotypes at various stages of development obtained from $Tnip1^{+/-}$ heterozygote intercrosses are shown. Total number of embryos/pups at each stage of development for each TNF genotype is indicated at right (total). The percentages of embryos that were homozygous deficient ($Tnip1^{-/-}$) among specific TNF genotypes and at specific stages of development are indicated in parentheses.

The failure to protect embryonic fetal livers from TNF-induced programmed cell death (PCD) prevents successful development of mice[3–6]. We thus interbred $Tnip1^{+/-}$ mice with $Tnip1^{-/-}$ mice to determine whether ABIN-1 supports embryonic development by regulating TNF-induced signals. Genotyping of live born mice from $Tnip1^{+/-}$ $Tnf^{+/-}$ intercrosses revealed that $Tnip1^{-/-}$ animals were obtained in roughly Mendelian ratios when animals were homozygous $Tnf^{-/-}$ mice (Table 1). Therefore, ABIN-1 sustains embryogenesis by regulating TNF signals, possibly by protecting cells against TNF-induced PCD.

To determine whether ABIN-1 directly protects cells from TNF-induced PCD, we tested the susceptibility to PCD of ABIN-1 deficient Jurkat T cells, HepG2 hepatoma cells and HT1080 fibrosarcoma cells as well as murine embryonic fibroblasts (MEFs). ABIN-1 deficient cells and $Tnip1^{-/-}$ MEFS were more sensitive to TNF-induced PCD in the presence or absence of CHX (cycloheximide) (Fig. 2a, Supplementary Fig. 2a–c). As ABIN-1 is expressed in unstimulated cells, we performed subsequent experiments in the presence of CHX to distinguish ABIN-1's anti-apoptotic functions from other translation dependent anti-apoptotic mechanisms. ABIN-1 deficient cells exhibited more caspase 8, caspase 3 and Bid cleavage when compared to control cells after treatment with TNF plus CHX (Fig. 2b and Supplementary Fig. 2d). PCD in $Tnip1^{-/-}$ cells was blocked by the caspase inhibitor ZVAD (benzyloxycarbonyl-valine-alanine-aspartate-fluoromethyl ketone) (Fig. 2c), confirming that increased PCD in these cells was caspase mediated. Finally, GFP–Flag–ABIN-1 protected $Tnip1^{-/-}$ cells from TNF-induced PCD and prevented caspase 8, Bid and caspase 3 cleavage (Fig. 2c, d). These results are consistent with a prior suggestion that ABIN-1 can protect hepatocytes from TNF-induced PCD[7]. Our experiments suggest that ABIN-1 directly protects multiple cell types against TNF-induced PCD, including both type I and type II cells[8].

Cells that exhibit deficient NF-κB signalling are hypersensitive to TNF-induced PCD[4–6]. NF-κB signalling was not decreased in $Tnip1^{-/-}$ cells compared to control cells after TNF treatment (Fig. 2e, Supplementary Fig. 3a–c). Levels of NF-κB dependent survival proteins such as Bcl-x_L, XIAP and A20 were expressed at normal levels in $Tnip1^{-/-}$ cells (Fig. 2e). Prolonged JNK signalling has also been associated with exaggerated PCD after TNF treatment[9–11]. However, $Tnip1^{-/-}$ MEFs and $Tnip1^{+/+}$ MEFS exhibited similar kinetics of phospho-JNK activation after TNF treatment, as well as normal p38 and ERK signalling (Fig. 2e). These experiments suggest that ABIN-1 directly protects cells against TNF-induced PCD, and that this function is not secondary to aberrant MAP kinase signalling.

[1]Department of Medicine, University of California at San Francisco, 513 Parnassus Avenue, S-1057, San Francisco, California 94143-0451, USA.
*These authors contributed equally to this work.

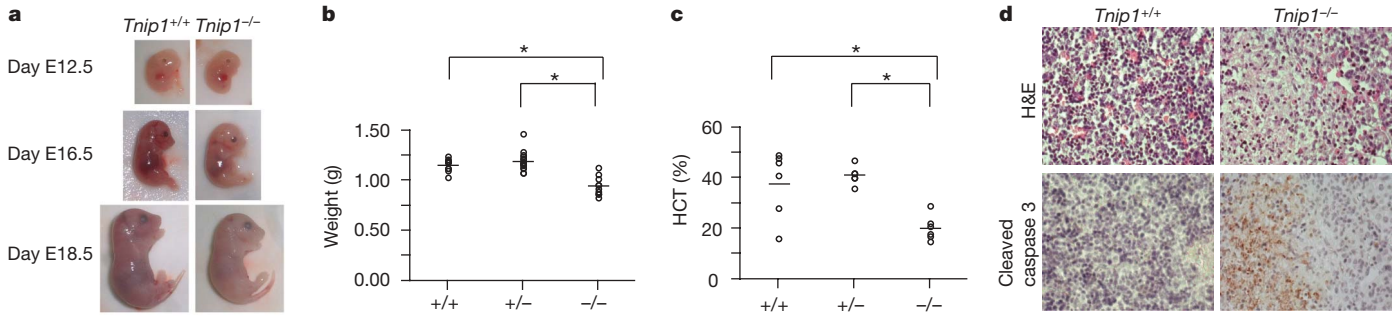


Figure 1 | ABIN-1 is required for embryonic development. **a**, Gross appearance of *Tnip1*$^{+/+}$ and *Tnip1*$^{-/-}$ embryos. **b**, Hypoplasia of E18.5 *Tnip1*$^{-/-}$ embryos. Circles, weights of individual embryos; horizontal bars, mean weights for each genotype. *Tnip1*$^{-/-}$ embryos weigh less than *Tnip*$^{+/+}$ and *Tnip1*$^{+/-}$ control embryos ($P < 0.01$ between *Tnip1*$^{+/+}$ and *Tnip1*$^{-/-}$ embryos; $P < 0.01$ between *Tnip1*$^{+/-}$ and *Tnip1*$^{-/-}$ embryos, $n = 35$). **c**, Anaemia of E18.5 *Tnip1*$^{-/-}$ embryos. *Tnip1*$^{-/-}$ embryos have markedly lower haematocrits (HCT) ($P < 0.02$ between *Tnip1*$^{+/+}$ and *Tnip1*$^{-/-}$ embryos; $P < 0.01$ between *Tnip1*$^{+/-}$ and *Tnip1*$^{-/-}$ embryos, $n = 18$). **d**, Haematoxylin and eosin (H&E) histology (upper panels) and cleaved caspase 3 immunohistochemistry (lower panels) of sequential sections from *Tnip1*$^{+/+}$ and *Tnip1*$^{-/-}$ fetal livers. Apoptotic patches were found in 5 of 7 *Tnip1*$^{-/-}$ fetal livers analysed and in none of 10 *Tnip1*$^{+/-}$ and *Tnip1*$^{+/+}$ fetal livers analysed ($P < 0.01$). *Statistically significant differences.

Prior studies suggested that ABIN-1 binds A20, a potent restrictor of TNF-induced NF-κB signalling, and that ABIN-1 inhibits NF-κB signalling[1,12]. *Tnip1*$^{-/-}$ cells exhibited slightly greater p-IκBα to IκBα ratios and IKKβ kinase activity than control cells, while *A20*$^{-/-}$ cells displayed significantly prolonged NF-κB signalling (Supplementary Fig. 3a, b). *Tnip1*$^{-/-}$ cells expressed similar levels of NF-κB specific DNA binding activity, and similar levels of NF-κB dependent messenger RNAs and proteins (Supplementary Fig. 3c–e). By contrast, *A20*$^{-/-}$ cells expressed significantly prolonged NF-κB DNA binding activity and markedly higher levels of most NF-κB target genes and proteins compared with either *Tnip1*$^{-/-}$ or control cells (Supplementary Fig. 3c–e). Hence, ABIN-1 deficiency reveals a subtle role for ABIN-1 in restricting proximate NF-κB signalling, which is associated with minimal effects on NF-κB dependent gene transcription. Meanwhile, A20 deficiency leads to markedly prolonged NF-κB signalling and increased NF-κB dependent gene transcription.

After TNFR engagement, death inducing signalling complexes (DISCs) including TRADD, FADD and caspase 8 may trigger caspase 8 cleavage and PCD[13,14]. ABIN-1 is expressed in resting cells, and ABIN-1 regulates TNF-induced PCD in the absence of protein

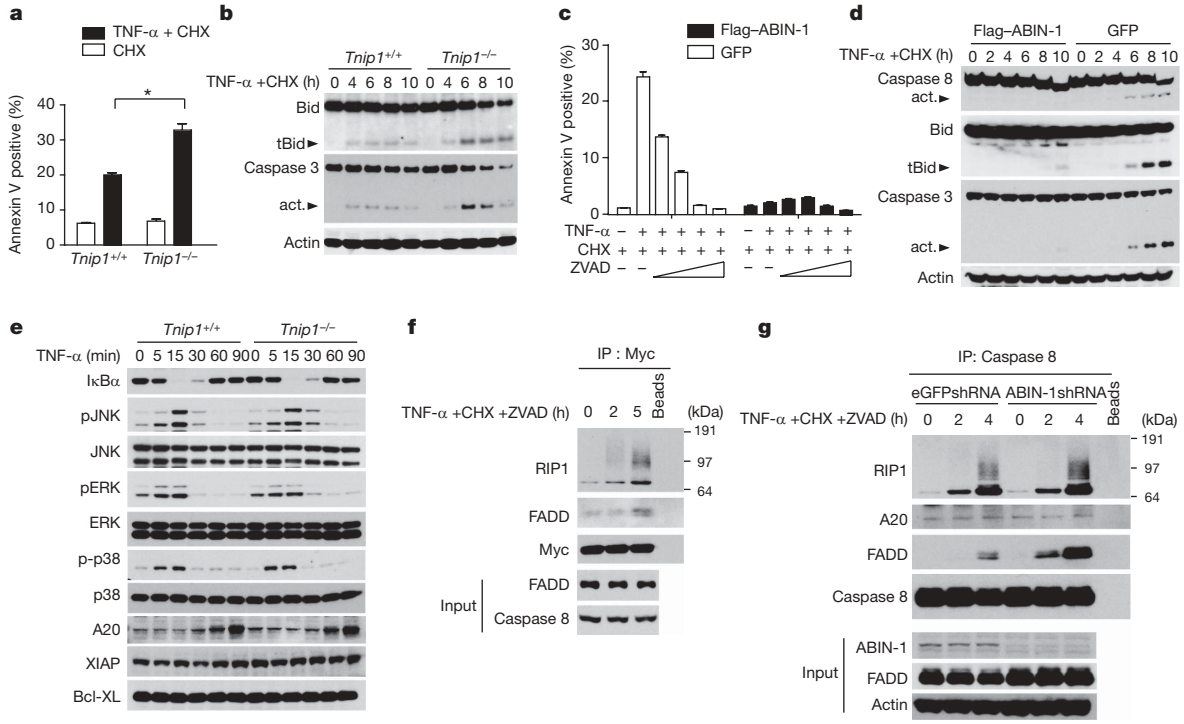


Figure 2 | ABIN-1 is required for protecting cells from TNF-induced PCD. **a**, TNF-induced PCD of *Tnip1*$^{+/+}$ and *Tnip1*$^{-/-}$ MEFs ($P < 0.01$ between *Tnip*$^{+/+}$ and *Tnip1*$^{-/-}$ cells treated with TNF plus CHX; means and standard deviations indicated, $n = 3$). **b**, TNF-induced caspase 3 and Bid expression in *Tnip1*$^{+/+}$ (left) and *Tnip1*$^{-/-}$ (right) MEFs. Cleaved, active forms of Bid (tBid) and caspase 3 (act.) indicated by arrows. Actin expression shown below as protein loading control. **c**, TNF-induced PCD of *Tnip1*$^{-/-}$ 3T3s expressing GFP and Flag–ABIN-1 (black columns on right) or GFP alone (white columns on left), after treatment with the indicated agents. Means and standard deviations indicated, $n = 3$. **d**, TNF-induced caspase 8, caspase 3 and Bid expression in *Tnip1*$^{-/-}$ 3T3s expressing GFP and Flag–ABIN-1 (left) or GFP alone (right). Cleaved, active forms of caspase 8, caspase 3 and Bid (tBid) indicated by arrows. Actin expression shown below as control. **e**, TNF-induced expression of MAP kinase signalling proteins and A20 in *Tnip1*$^{-/-}$ and control MEFs. **f**, TNF-induced recruitment of RIP1 and FADD to Myc (ABIN-1). Expression of Myc–ABIN-1 in IPs, and expression of FADD and caspase 8 in whole cell lysates ('input') shown below as controls. **g**, TNF-induced recruitment of RIP1, FADD and A20 to caspase 8 in ABIN-1 deficient and control HT1080 cells. Expression of caspase 8 in IPs and expression of ABIN-1, FADD and actin in whole cell lysates ('input') shown below as controls. All data are representative of 3–5 independent experiments. *Statistically significant differences.

synthesis. We thus hypothesized that ABIN-1 might directly regulate TNF-induced DISCs. We found that ABIN-1 indeed interacts with RIP1 and FADD under conditions that induce PCD (Fig. 2f). To determine whether ABIN-1 directly regulates association of DISC proteins, we stimulated ABIN-1 deficient and control cells with TNF, CHX and ZVAD, then immunoprecipitated endogenous caspase 8, and finally examined the association of DISC proteins. Endogenous RIP-1 was recruited normally to caspase 8 under these conditions (Fig. 2g). By contrast, the amount of endogenous FADD associated with caspase 8 was significantly greater in ABIN-1 deficient cells, indicating that ABIN-1 is essential for inhibiting the interaction of caspase 8 with FADD (Fig. 2g).

To further investigate how ABIN-1 may regulate TNF-induced PCD, we noted that a peptide sequence overlapping the ABIN homology domain 2 (AHD2) of ABIN-1 is homologous to a larger, recently defined domain within IKKγ/NEMO that binds polyubiquitin chains and ubiquitylated RIP, called the NEMO ubiquitin binding (NUB) domain[15–17] (Fig. 3a). In addition, ABIN-1 has a leucine zipper upstream of this domain (Fig. 3a). Hypothesizing that ABIN-1, like IKKγ, may be a ubiquitin binding protein, we tested the ability of recombinant ABIN-1 proteins to bind His tagged ubiquitin chains. These studies showed that GST–ABIN-1 binds to His tagged polyubiquitin chains in a GST pull-down assay (Fig. 3b). GST–ABIN-1 binds both K48-linked as well as K63-linked polyubiquitin chains, and displays a preference for chains that are at least three ubiquitin moieties in length (Fig. 3b). GST–ABIN-1 did not bind ubiquitin monomers (Fig. 3b). These findings are consistent with a recent study demonstrating binding of ABIN-1 to ubiquitin chains[18].

To further establish and localize the ubiquitin binding domain of ABIN-1, carboxy-terminal and amino-terminal truncation mutants of ABIN-1 were tested for ubiquitin binding. C-terminal GST–ABIN-1, which contains AHD2, binds K63 linked poly-ubiquitin chains while an N-terminal GST–ABIN-1 protein lacking this domain does not (Fig. 3c). Several residues within the NUB domain of IKKγ that are required for ubiquitin binding are conserved in ABIN-1. C-terminal ABIN-1 proteins bearing either QQ (glutamine) to EE (glutamic acid) substitutions at positions 477/478 ('QQ477EE'), or a F (phenylalanine) to S (serine) mutation at position 482 ('F482S'), failed to bind ubiquitin chains (Fig. 3d). Thus, ABIN-1 is a ubiquitin sensing protein that utilizes a NUB-like domain to bind polyubiquitin chains.

To determine whether ABIN-1's ubiquitin binding activity mediates its anti-apoptotic function, we introduced wild-type or mutant forms of Flag tagged ABIN-1 into *Tnip1*$^{-/-}$ 3T3 cells using a GFP expressing virus, purified productively infected cells by FACS sorting, and tested

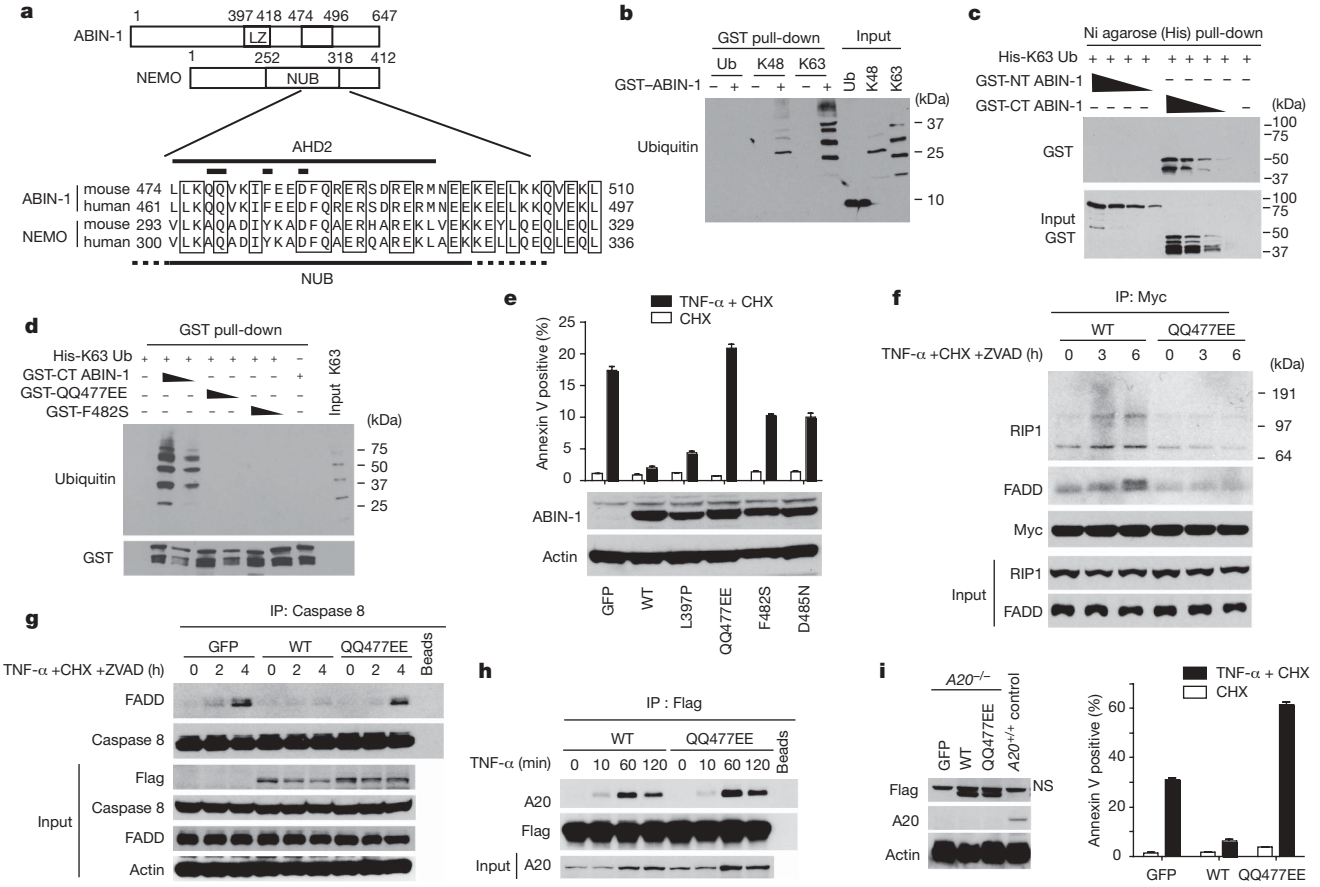


Figure 3 | ABIN-1 is a ubiquitin sensor that uses a NUB domain to bind to DISC signalling complexes and protect cells from TNF-induced PCD. **a**, Sequence alignment of AHD2 domain of ABIN-1 with the NEMO ubiquitin binding (NUB) domain of NEMO/IKKγ. Thick bars mark residues mutated in QQ477EE, F482S and D485N ABIN-1 mutants. **b**, Ubiquitin binding to GST–ABIN-1 in GST pull-down assay. Ubiquitin bound to GST–ABIN-1 shown in left six lanes; input ubiquitin chains are shown at right. **c**, GST–ABIN-1 proteins bound to K63-linked His-polyubiquitin chains in nickel agarose pull-down assay (upper panel). Input GST–ABIN-1 proteins shown below. **d**, 6×His-K63 linked ubiquitin chains bound to GST-C-terminal ABIN-1 proteins (upper panel). Input GST–ABIN-1 proteins shown below. **e**, TNF-induced PCD of *Tnip1*$^{-/-}$ 3T3s complemented with the indicated Flag–ABIN-1 proteins or GFP alone. Flag–ABIN-1 and actin expression levels shown below. **f**, Recruitment of RIP1 or FADD to wild-type (WT) or mutant Myc–ABIN-1. Myc–ABIN-1 expression in anti-Myc IPs and RIP1 and FADD expression in whole cell lysates ('input') are shown. **g**, Recruitment of FADD to caspase 8 in presence of WT or QQ477EE mutant Flag–ABIN-1, or no ABIN-1 ('GFP'). Expression of caspase 8 in IPs and expression of Flag–ABIN-1, caspase 8, FADD and actin in whole cell lysates ('input') are shown. **h**, Recruitment of A20 to WT or mutant ('QQ477EE') Flag–ABIN-1 proteins. Expression of Flag–ABIN-1 in IPs and A20 in whole cell lysates ('input') shown. **i**, Anti-apoptotic activity of ABIN-1 in *A20*$^{-/-}$ 3T3s expressing no ABIN-1 ('GFP'), WT ABIN-1 ('WT'), or mutant ABIN-1 ('QQ477EE'). Means and standard deviations indicated, $n = 3$. Expression of Flag–ABIN-1, A20 and actin proteins shown at left (NS, non-specific band above Flag–ABIN-1 band). Data are representative of 3–5 independent experiments.

the cells for susceptibility to TNF-induced PCD. These experiments revealed that wild-type ABIN-1 and a leucine zipper mutant L397P protected *Tnip1*$^{-/-}$ cells from TNF-induced PCD (Fig. 3e). By contrast, ubiquitin binding domain mutants of ABIN-1 (QQ477EE, F482S and D485N) failed to fully protect *Tnip1*$^{-/-}$ cells from TNF-induced PCD (Fig. 3e). To determine whether ABIN-1's ubiquitin binding activity mediates ABIN-1's capacity for regulating DISC formation, we tested whether a ubiquitin binding deficient form of ABIN-1 binds to this complex and whether it regulates the association of DISC proteins. Wild-type, but not QQ477EE mutant, ABIN-1 was recruited to RIP1 and FADD (Fig. 3f). Moreover, wild-type, but not QQ477EE mutant, ABIN-1 blocked the recruitment of endogenous FADD to caspase 8 (Fig. 3g). Therefore, ABIN-1 requires its ubiquitin binding domain to interact with DISC proteins, to inhibit FADD–caspase 8 binding, and to protect cells from TNF-induced PCD.

ABIN-1 binds to the ubiquitin editing enzyme A20, which also protects cells against TNF-induced PCD, so we investigated whether ABIN-1 might block TNF-induced PCD in an A20 dependent fashion[1,12]. A20 protein was induced by TNF in both *Tnip1*$^{+/+}$ and *Tnip1*$^{-/-}$ MEFs, so ABIN-1 is not required for A20 expression (Fig. 2e). A20 association with caspase 8 does not depend upon ABIN-1, arguing against a critical role for ABIN-1 in recruiting A20 to the DISC (Fig. 2g). A20 was recruited normally to both wild-type and QQ477EE forms of ABIN-1, indicating that ABIN-1 does not require ubiquitin binding to bind A20 (Fig. 3h). Finally, wild-type, but not QQ477EE mutant, ABIN-1 blocks TNF-induced PCD in *A20*$^{-/-}$ cells, showing that ABIN-1 does not require A20 to perform its anti-apoptotic function (Fig. 3i).

In summary, we have demonstrated that ABIN-1 is a novel ubiquitin-sensing protein that inhibits FADD–caspase 8 association in the DISC, protects cells against TNF-induced PCD, and sustains embryonic development. The recent unveiling of non-degradative ubiquitin modifications has led to the discovery of ubiquitin sensors that propagate NF-κB signals[16,17,19–21]. Much less is known about ubiquitin sensors and PCD signalling[22,23]. Our current studies provide evidence that ubiquitin sensors also restrict signal transduction pathways, and that they regulate PCD signalling. These findings highlight the central roles of ubiquitin sensors in regulating physiological cell survival decisions *in vivo*.

METHODS SUMMARY

Generation of ABIN-1 deficient (*Tnip1*$^{-/-}$) mice. The generation of a gene targeting construct, targeted embryonic stem (ES) cells, and *Tnip1* deficient mice are described in Methods and Supplementary Fig. 1 legend.

Lentiviral shRNA knock-down of ABIN-1 in human cell lines. Lentiviruses bearing shRNA sequences specific to ABIN-1 were purchased from Open Biosystems and used to infect Jurkat, HepG2 or HT1080 cells.

Cell-free ubiquitin binding assays. Poly-histidine-tagged ubiquitin and polyubiquitin chains were purchased from Boston Biochem. Recombinant GST–ABIN-1 proteins were purified from *Escherichia coli* BL21 bacteria. Ubiquitin binding assays were performed using either glutathione agarose beads or nickel agarose beads for precipitating ABIN-1–ubiquitin complexes.

Lentiviral transduction of ABIN-1 expression constructs. Lentiviral expression vectors expressing GFP and Flag-tagged ABIN were used to infect cells. In some experiments, infected GFP-positive cells were purified by cell sorting before use in cell signalling and cell death assays.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 26 September 2007; accepted 24 October 2008.
Published online 7 December 2008.

1. Heyninck, K. *et al.* The zinc finger protein A20 inhibits TNF-induced NF-κB-dependent gene expression by interfering with an RIP- or TRAF2-mediated transactivation signal and directly binds to a novel NF-κB-inhibiting protein ABIN-1. *J. Cell Biol.* **145,** 1471–1482 (1999).
2. Opipari, A. W. Jr, Boguski, M. S. & Dixit, V. M. The A20 cDNA induced by tumor necrosis factor α encodes a novel type of zinc finger protein. *J. Biol. Chem.* **265,** 14705–14708 (1990).
3. Doi, T. S. *et al.* Absence of tumor necrosis factor rescues RelA-deficient mice from embryonic lethality. *Proc. Natl Acad. Sci. USA* **96,** 2994–2999 (1999).
4. Beg, A. A. & Baltimore, D. An essential role for NF-κB in preventing TNF-α-induced cell death. *Science* **274,** 782–784 (1996).
5. Wang, C. Y., Mayo, M. W. & Baldwin, A. S. Jr. TNF- and cancer therapy-induced apoptosis: Potentiation by inhibition of NF-κB. *Science* **274,** 784–787 (1996).
6. Van Antwerp, D. J., Martin, S. J., Kafri, T., Green, D. R. & Verma, I. M. Suppression of TNF-α-induced apoptosis by NF-κB. *Science* **274,** 787–789 (1996).
7. Wullaert, A. *et al.* Adenoviral gene transfer of ABIN-1 protects mice from TNF/galactosamine-induced acute liver failure and lethality. *Hepatology* **42,** 381–389 (2005).
8. Barnhart, B. C., Alappat, E. C. & Peter, M. E. The CD95 type I/type II model. *Semin. Immunol.* **15,** 185–193 (2003).
9. De Smaele, E. *et al.* Induction of gadd45β by NF-κB downregulates pro-apoptotic JNK signalling. *Nature* **414,** 308–313 (2001).
10. Tang, G. *et al.* Inhibition of JNK activation through NF-κB target genes. *Nature* **414,** 313–317 (2001).
11. Chang, L. *et al.* The E3 ubiquitin ligase itch couples JNK activation to TNFα-induced cell death by inducing c-FLIP(L) turnover. *Cell* **124,** 601–613 (2006).
12. Lee, E. G. *et al.* Failure to regulate TNF-induced NF-κB and cell death responses in A20-deficient mice. *Science* **289,** 2350–2354 (2000).
13. Micheau, O. & Tschopp, J. Induction of TNF receptor I-mediated apoptosis via two sequential signaling complexes. *Cell* **114,** 181–190 (2003).
14. Schneider-Brachert, W. *et al.* Compartmentalization of TNF receptor 1 signaling: Internalized TNF receptosomes as death signaling vesicles. *Immunity* **21,** 415–428 (2004).
15. Heyninck, K., Kreike, M. M. & Beyaert, R. Structure-function analysis of the A20-binding inhibitor of NF-κ B activation, ABIN-1. *FEBS Lett.* **536,** 135–140 (2003).
16. Ea, C. K., Deng, L., Xia, Z. P., Pineda, G. & Chen, Z. J. Activation of IKK by TNFα requires site-specific ubiquitination of RIP1 and polyubiquitin binding by NEMO. *Mol. Cell* **22,** 245–257 (2006).
17. Wu, C. J., Conze, D. B., Li, T., Srinivasula, S. M. & Ashwell, J. D. Sensing of Lys 63-linked polyubiquitination by NEMO is a key event in NF-κB activation. *Nature Cell Biol.* **8,** 398–406 (2006).
18. Wagner, S. *et al.* Ubiquitin binding mediates the NF-κB inhibitory potential of ABIN proteins. *Oncogene* **27,** 3739–3745 (2008).
19. Deng, L. *et al.* Activation of the IκB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. *Cell* **103,** 351–361 (2000).
20. Wang, C. *et al.* TAK1 is a ubiquitin-dependent kinase of MKK and IKK. *Nature* **412,** 346–351 (2001).
21. Chen, Z. J. Ubiquitin signalling in the NF-κB pathway. *Nature Cell Biol.* **7,** 758–765 (2005).
22. Wertz, I. E. & Dixit, V. M. Ubiquitin-mediated regulation of TNFR1 signaling. *Cytokine Growth Factor Rev.* **19,** 313–324 (2008).
23. Lee, J. C. & Peter, M. E. Regulation of apoptosis by ubiquitination. *Immunol. Rev.* **193,** 39–47 (2003).
24. Boone, D. L. *et al.* The ubiquitin-modifying enzyme A20 is required for termination of Toll-like receptor responses. *Nature Immunol.* **5,** 1052–1060 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the NIH (A.M.), UCSF Liver Center Pathology and Flow Cytometry Facilities, UCSF Transgenic and Targeted Mutagenesis Core Facility, a fellowship from the Crohn's and Colitis Foundation of America (S.O.), a pre-doctoral NSF fellowship (J.A.C.) and the Rainin Foundation.

Author Contributions S.O. and E.E.T. contributed equally to this work. S.O. and J.A.C. performed the cell signalling experiments. J.A.C., E.E.T. and S.O. analysed *Tnip1*$^{-/-}$ embryos. S.O. performed DISC analyses. E.E.T. performed the ubiquitin binding experiments. E.E.T. and S.C. designed and generated the ABIN-1 gene targeting construct. B.A.M. generated Tnip1$^{-/-}$ *ES cells and mice with assistance from S.C. R.A. genotyped and maintained* Tnip1$^{-/-}$ *mice and performed timed matings. J.B., N.S., B.L. and T.W. performed mutagenesis and generated expression constructs. B.Y. analysed histological and immunohistochemical studies of* Tnip1$^{-/-}$ *embryos. A.M. and B.A.M. supervised the overall project, and wrote and edited the manuscript.*

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.M. (averil.ma@ucsf.edu).

METHODS

Generation of ABIN-1 deficient (*Tnip1*$^{-/-}$) mice. Recombineering was used to generate a gene targeting construct from a bacterial artificial chromosome (BAC) containing the *Tnip1* gene (C57BL/6J inbred strain). C57BL/6J inbred PRX-B6T ES cells were transfected with this construct, and successfully targeted ES cells were identified by Southern blot analysis. Blastocyst injections of targeted ES cells were performed by the UCSF Transgenic Core. Mice bearing this targeted allele in the germline were interbred with E2a-Cre transgenic mice to delete intervening sequences including exons 12–15 and generate the null allele. These mice were then bred to B6 mice to eliminate the E2a-Cre transgene and generate *Tnip1*$^{+/-}$ mice.

Programmed cell death assays. ABIN-1 competent and deficient cells were treated with TNF (10 ng ml^{-1}) plus cycloheximide (10 μg ml^{-1}) and assayed for programmed cell death by flow cytometric quantitation of Annexin V reactive cells using an LSRII flow cytometer and Flo-Jo software. Cell death signalling in these cells was assessed by preparing whole cell lysates in lysis buffer (50 mM HEPES, 120 mM NaCl, 1 mM EDTA, 0.1% NP-40) and immunoblotting for expression and cleavage of cell death proteins using commercial antibodies (caspase 3, Cell Signaling; caspase 8, BD Pharmingen, Alexis Biochemicals; BID, Santa Cruz). In some assays, Z-VAD-FMK was used (Alexis Biochemicals).

Cell signalling assays. Activation of cell signalling cascades was assessed by stimulating cells with the indicated ligands (for example, TNF), preparing whole cell lysates in lysis buffer (above), and immunoblotting for the expression of various proteins using commercial antibodies (phospho-IκBα, IκBα, phospho-MAPK, MAPK, phospho-JNK, JNK, phospho-p38, p-38, XIAP (Cell Signaling), Bcl-x (BD Pharmingen), ABIN-1 (Zymed), actin (Calbiochem)). A20 protein was detected using either a commercial anti-human A20 antibody (eBioscience), or a polyclonal anti-murine A20 antibody[24]. Recruitment of co-associated signalling proteins was assayed by lysing cells in lysis buffer (10 mM Tris, 140 mM NaCl, 0.2% NP-40, 10% glycerol and protease inhibitors), immunoprecipitating Flag-tagged ABIN-1 protein with anti-Flag beads (Sigma), Myc-tagged ABIN-1 with anti-c-Myc antibody (Santa Cruz Biotechnology), or caspase 8 with anti-caspase 8 (Santa Cruz Biotechnology) and immunoblotting co-associated FADD, caspase 8, A20, or RIP using commercial antibodies (RIP (BD Transduction Labs, Santa Cruz Biotechnology), FADD (Assay Designs)). IKK kinase assays were performed by immunoprecipitating lysates with anti-IKKγ antibody (Santa Cruz Biotechnology), incubating with recombinant GST–IκBα, and measuring phospho-IκBα by immunoblotting. Control for immunoprecipitation efficiency was performed with anti-IKKγ antibody (Becton Dickinson). To perform NF-κB EMSAs, 10 μg of protein from MEF nuclear lysates were incubated with a NF-κB consensus oligonucleotide (Santa Cruz Biotechnology), separated on 6% acrylamide gels, and developed with a chemiluminescent detection kit (Pierce).

Real-time PCR assays of mRNA expression. Real-time PCR analyses were performed using TaqMan Gene Expression kit (Applied Biosystems), according to manufacturer's instructions.

Lentiviral shRNA knock-down of ABIN-1 in human cell lines. Lentiviral constructs bearing shRNA sequences specific to ABIN-1 were purchased from Open Biosystems. Constructs were co-transfected with expression constructs for helper virus R89.1, VSV-G and tat into 293T cells, and replication deficient virus from the resulting supernatants was used to infect Jurkat, HepG2 and HT1080 cells. The degree of ABIN-1 protein reduction was assessed by immunoblotting of whole cell lysates. Two of five shRNA sequences (TRCN8681 ('shRNA 81') and TRCN8685 ('shRNA 85')) that caused the greatest reduction of ABIN-1 protein expression were selected for cell signalling and cell death studies. In some experiments, these two shRNA-containing viruses were combined in a single infection. Control cells included cells infected with lentiviruses devoid of shRNA sequences, or expressing eGFP specific sequences.

Cell-free ubiquitin binding assays. Poly-histidine-tagged ubiquitin and poly-ubiquitin chains were purchased from Boston Biochem. Wild-type and mutant ABIN-1 cDNAs were generated by site-directed mutatgenesis (Quick Change Kit, Stratagene), after which they were subcloned into bacterial expression vectors. Recombinant GST–ABIN-1 proteins were purified from bacteria using glutathione agarose beads. Ubiquitin-binding assays with GST–ABIN-1 proteins were performed either by (1) incubating glutathione agarose beads bound to GST–ABIN-1 proteins with free poly-histidine-ubiquitin for 10 min at room temperature in ubiquitin binding buffer (25 mM HEPES, 150 mM KCl, 2 mM $MgCl_2$, 0.5% Triton X-100, 1 mM EGTA, and 1 mg ml^{-1} BSA), followed by serial washes with binding buffer and precipitation of glutathione agarose beads by centrifugation, or (2) by incubating free GST–ABIN-1 proteins with nickel agarose beads bound to poly-histidine-ubiquitin in ubiquitin binding buffer supplemented with 5 mM imidazole, followed by serial washes and centrifugation of nickel agarose beads. In both cases, bead bound ABIN-1–ubiquitin complexes were analysed by immunoblotting for GST (ABIN-1 protein) and/or ubiquitin using commercial antibodies (GST (Cell Signaling), ubiquitin (Santa Cruz Biotechnology)).

Expression of ABIN-1 proteins. Lentiviral expression vectors expressing GFP and Flag-tagged ABIN-1 constructs were generated from a lentiviral pWPI construct. These constructs were co-transfected along with helper constructs into 293T cells to generate infectious lentivirus, which was then used to infect various cell types. In some assays, productively infected cells were purified by flow cytometry for GFP expressing cells before use in cell signalling and cell death assays. In some experiments, transient expression of Myc–ABIN-1 was performed in 293T cells.

Vol 457|12 February 2009|doi:10.1038/nature07762

LETTERS

Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression

Arun Sreekumar[1,2,3,4], Laila M. Poisson[5]*, Thekkelnaycke M. Rajendiran[1,3]*, Amjad P. Khan[1,3]*, Qi Cao[1,3], Jindan Yu[1,3], Bharathi Laxman[1,3], Rohit Mehra[1,3], Robert J. Lonigro[1,4], Yong Li[1,3], Mukesh K. Nyati[4,6], Aarif Ahsan[6], Shanker Kalyana-Sundaram[1,3], Bo Han[1,3], Xuhong Cao[1,3], Jaeman Byun[7], Gilbert S. Omenn[2,7,8], Debashis Ghosh[4,5,11], Subramaniam Pennathur[2,4,7], Danny C. Alexander[12], Alvin Berger[12], Jeffrey R. Shuster[12], John T. Wei[4,9], Sooryanarayana Varambally[1,3,4], Christopher Beecher[1,2,3] & Arul M. Chinnaiyan[1,2,3,4,9,10]

Multiple, complex molecular events characterize cancer development and progression[1,2]. Deciphering the molecular networks that distinguish organ-confined disease from metastatic disease may lead to the identification of critical biomarkers for cancer invasion and disease aggressiveness. Although gene and protein expression have been extensively profiled in human tumours, little is known about the global metabolomic alterations that characterize neoplastic progression. Using a combination of high-throughput liquid-and-gas-chromatography-based mass spectrometry, we profiled more than 1,126 metabolites across 262 clinical samples related to prostate cancer (42 tissues and 110 each of urine and plasma). These unbiased metabolomic profiles were able to distinguish benign prostate, clinically localized prostate cancer and metastatic disease. Sarcosine, an *N*-methyl derivative of the amino acid glycine, was identified as a differential metabolite that was highly increased during prostate cancer progression to metastasis and can be detected non-invasively in urine. Sarcosine levels were also increased in invasive prostate cancer cell lines relative to benign prostate epithelial cells. Knockdown of glycine-*N*-methyl transferase, the enzyme that generates sarcosine from glycine, attenuated prostate cancer invasion. Addition of exogenous sarcosine or knockdown of the enzyme that leads to sarcosine degradation, sarcosine dehydrogenase, induced an invasive phenotype in benign prostate epithelial cells. Androgen receptor and the ERG gene fusion product coordinately regulate components of the sarcosine pathway. Here, by profiling the metabolomic alterations of prostate cancer progression, we reveal sarcosine as a potentially important metabolic intermediary of cancer cell invasion and aggressivity.

To profile the 'metabolome' during prostate cancer progression, we used both liquid and gas chromatography coupled with mass spectrometry[3] to interrogate the relative levels of metabolites across 262 prostate-related biospecimens (outlined in Supplementary Fig. 1). Specifically, 42 tissue samples and 110 matched specimens of plasma and post-digital-rectal-exam urine from biopsy-positive cancer patients ($n = 59$) and biopsy-negative control individuals ($n = 51$) were assayed (Fig. 1a). A total of 1,126 metabolites were quantified and, as expected, only a small percentage of these metabolites (15.6%) were shared across the disparate biospecimen types (Fig. 1a).

Evaluation of the unbiased metabolomic profiles of plasma or urine did not identify robust differences between biopsy-positive and -negative individuals. For plasma, 20 out of 478 (4%) metabolites were differential (Wilcoxon $P < 0.05$), with a false discovery rate (FDR) of 99%. Likewise, for urine, 36 out of 583 (6%) metabolites were differential (Wilcoxon $P < 0.05$), with an FDR of 67%. Thus, our initial focus was directed towards understanding the tissue metabolomic profiles as they showed more robust alterations.

Tissue samples were derived from benign adjacent prostate ($n = 16$), clinically localized prostate cancer ($n = 12$, PCA) and metastatic prostate cancer ($n = 14$) patients. Selection of metastatic tissue samples from different sites (see Supplementary Table 2) minimized characterization of analytes specific to cells of non-prostatic origin. In total, high-throughput profiling of the tissues quantitatively detected 626 metabolites (175 named, 19 isobars and 432 metabolites without identification), of which 82.3% (515 out of 626) were shared by the three diagnostic classes (Fig. 1b). Notably, there were 60 metabolites found in PCA and/or metastatic tumours but not in benign prostate. These profiles were displayed as a heat map (Fig. 1c) and *z*-score plot (Fig. 1d). In the latter, benign-based *z*-scores were plotted for each of the 626 metabolites. The plots revealed robust metabolic alterations in metastatic tumours (*z*-score range: −13.6 to 81.9) compared to fewer changes in clinically localized prostate cancer samples (*z*-score range: −7.7 to 45.8).

We identified the differential metabolites between the PCA and benign samples using a two-sided Wilcoxon rank-sum test coupled with a permutation test ($n = 1{,}000$). A total of 87 out of 518 metabolites were differential across these two classes ($P < 0.05$, corresponding to a 23% FDR). For visualizing the relationship between the 87 altered metabolites, hierarchical clustering was used to arrange the metabolites on the basis of their relative levels across samples (Fig. 2a). Among the perturbed metabolites, 50 were increased in PCA samples whereas 37 were downregulated. Figure 2b displays the relative levels of the 37 named metabolites that were differential between benign prostate and PCA samples. Similarly, 124 out of 518 metabolites were found to be increased in the metastatic samples compared to the localized tumours and 102 compounds were downregulated ($P < 0.05$, corresponding to a 4% FDR). Figure 2c displays the levels of the 91 named metabolites altered in metastatic samples. A subset of six metabolites including sarcosine, uracil, kynurenine, glycerol-3-phosphate, leucine and proline were significantly increased on disease progression from benign to PCA to metastatic prostate cancer. These metabolites could potentially serve as biomarkers for progressive disease, one of the factors that motivated us to examine sarcosine in greater detail.

[1]The Michigan Center for Translational Pathology, [2]Center for Computational Medicine and Biology, [3]Department of Pathology, [4]The Comprehensive Cancer Center, [5]Department of Biostatistics, [6]Department of Radiation Oncology, [7]Department of Internal Medicine, [8]Department of Human Genetics, [9]Department of Urology, [10]Howard Hughes Medical Institute, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA. [11]Department of Statistics and Huck Institute of Life Sciences, Penn State University, Pennsylvania 16802, USA. [12]Metabolon, Inc. 800 Capitola Drive, Suite 1 Durham, North Carolina 27713, USA.
*These authors contributed equally to this work.

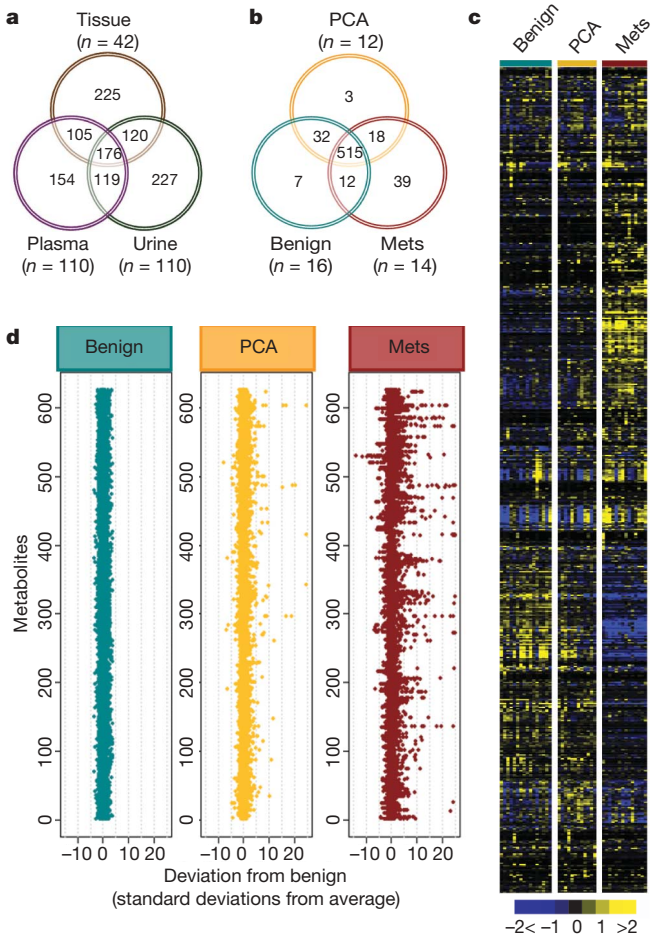


Figure 1 | Metabolomic profiling of prostate cancer. a, Venn diagram of the total metabolites detected across 42 prostate-related tissues and 110 matched plasma and urine samples. **b**, Venn diagram of 626 metabolites in tissues measured across 16 benign adjacent prostate tissues, 12 clinically localized prostate cancers (PCA) and 14 metastatic prostate cancers (Mets). **c**, Heat map representation of unsupervised hierarchical clustering of the data in **b** (rows) grouped by sample type (columns). Shades of yellow and blue represent an increase and decrease of a metabolite, respectively, relative to the median metabolite levels (see colour scale). **d**, *z*-score plots for the data in **b** normalized to the mean of the benign prostate samples (truncated at 25 s.d. for clarity, see Supplementary Methods for procedural details).

Mapping the differential metabolomic profiles to their respective biochemical pathways as outlined in the Kyoto Encyclopedia of Genes and Genomes (KEGG, release 41.1, http://www.genome.jp/kegg, Supplementary Fig. 8) revealed an increase in amino acid metabolism and nitrogen breakdown pathways during cancer progression to metastatic disease. A similar enrichment network of amino acid metabolism was also identified by the bioinformatics tool Oncomine Concept Map[4,5](OCM, http://www.oncomine.org, $P = 6 \times 10^{-13}$, Supplementary Figs 9 and 10a), supporting our earlier gene-expression-based prediction of androgen-induced protein synthesis as an early event during prostate cancer development[5]. Additionally, OCM found strong enrichment for increased 'methyltransferase activity' (Supplementary Fig. 10b, $P = 7.7 \times 10^{-8}$) among metabolites upregulated in metastatic samples. This corroborates previous studies from our group and others showing an increase of the histone methyltransferase EZH2 in metastatic tumours[6–11].

Because amino acid metabolism and methylation were enriched during prostate cancer progression, we focused on differential metabolites that characterize these processes and additionally show a progressive increase from benign to PCA to metastatic disease. The amino acid metabolite sarcosine, an *N*-methyl derivative of glycine, fits these criteria. Notably, metastatic samples showed markedly increased levels of sarcosine in 79% of the specimens analysed (chi-squared test, $P = 0.0538$), whereas 42% of the PCA samples showed an increase in the levels of this metabolite (Fig. 2a–c). Importantly, none of the benign samples had detectable levels of sarcosine. Taken together, this indicated the possible utility of sarcosine in monitoring disease progression and aggressiveness.

To confirm this pattern of sarcosine increase in cancer progression, we developed a highly sensitive and specific isotope dilution gas chromatography–mass spectrometry (GC–MS) method for accurately quantifying the metabolite from biospecimens (limit of detection = 10 femtomoles, Supplementary Fig. 11). In an independent set of 89 tissue samples (Supplementary Table 6), sarcosine levels were significantly increased in PCA specimens ($n = 36$) compared to benign adjacent prostate samples ($n = 25$, Wilcoxon $P = 4.34 \times 10^{-11}$, Fig. 3a). Additionally, there was an even greater increase of sarcosine in metastatic samples ($n = 28$) compared to organ-confined disease (Wilcoxon $P = 6.02 \times 10^{-11}$, Fig. 3a). In contrast, sarcosine was undetectable in adjacent non-neoplastic tissues from patients with metastatic disease (Supplementary Fig. 12a–c).

These findings led us to explore the potential of sarcosine as a candidate for future development in biomarker panels for early disease detection and aggressivity prediction. Towards this end, we monitored its levels in urine specimens from biopsy-positive and -negative individuals, most of whom have increased levels of prostate-specific antigen (PSA) (>4.0 ng ml^{-1}) and in which prostate needle biopsy was used for diagnosis. This is a particularly challenging cohort as, in addition to these men being at high risk for prostate cancer, even a negative needle biopsy does not rule out the presence of cancer due to sampling issues. Sarcosine was found to be significantly higher in urine sediments ($n = 49$, Wilcoxon $P = 0.0004$, Fig. 3b) and supernatants ($n = 59$, Wilcoxon $P = 0.0025$, Supplementary Fig. 14a) derived from biopsy-positive prostate cancer patients compared to biopsy-negative controls ($n = 44$ and $n = 51$, respectively; Supplementary Tables 7 and 8). The overall receiver operating characteristic curves for sarcosine indicate that its predictive value is modest, with an area under the curve (AUC) of 0.71 for urine sediments and 0.67 for supernatants (Supplementary Fig. 14b, c). Notably, an AUC of 1.0 indicates perfect prediction and an AUC of 0.5 indicates prediction equivalent to random selection. When restricted to samples having PSA in the clinical grey zone of 2–10 ng ml^{-1} ($n = 53$), sarcosine performed better than PSA in delineating the two diagnostic classes with an AUC of 0.69 (95% CI: 0.55, 0.84) compared to an AUC of 0.53 (95% CI: 0.37, 0.69) for PSA (Supplementary Fig. 15). Thus, sarcosine may have the potential to identify patients with modestly increased PSA that are likely to have a positive prostate biopsy.

To determine whether the sarcosine increase in prostate cancer has biological relevance, we measured its levels in prostate cancer cell lines VCaP, DU145, 22RV1 and LNCaP ($n = 3$ each) as well as in their benign epithelial counterparts, primary benign prostate epithelial cells (PrEC, $n = 2$) and immortalized benign RWPE prostate cells ($n = 3$). Significantly increased levels of the metabolite were found in prostate cancer cells compared to the benign cells (analysis of variance, ANOVA, $P = 0.0218$, Fig. 3c). Additionally, sarcosine levels correlated well with cell invasiveness (Spearman's correlation coefficient: 0.943, $P = 0.0048$). On the basis of our earlier findings that EZH2 overexpression in benign cells could mediate cell invasion and neoplastic progression[6,10,11], sarcosine levels were assessed on modulation of EZH2 expression. Interestingly, overexpression of EZH2 in benign prostate epithelial cells increased sarcosine levels (4.5-fold, Supplementary Fig. 16a), whereas its knockdown in DU145 prostate cancer cells diminished the levels of the metabolite (Supplementary Fig. 16a–c). To determine whether sarcosine has a more direct role in this process, we added the metabolite to non-invasive benign prostate epithelial cells. Alanine, an isomer of sarcosine, was used as a control for these experiments. Remarkably, the mere addition of exogenous sarcosine imparted an invasive phenotype to benign prostate epithelial cells

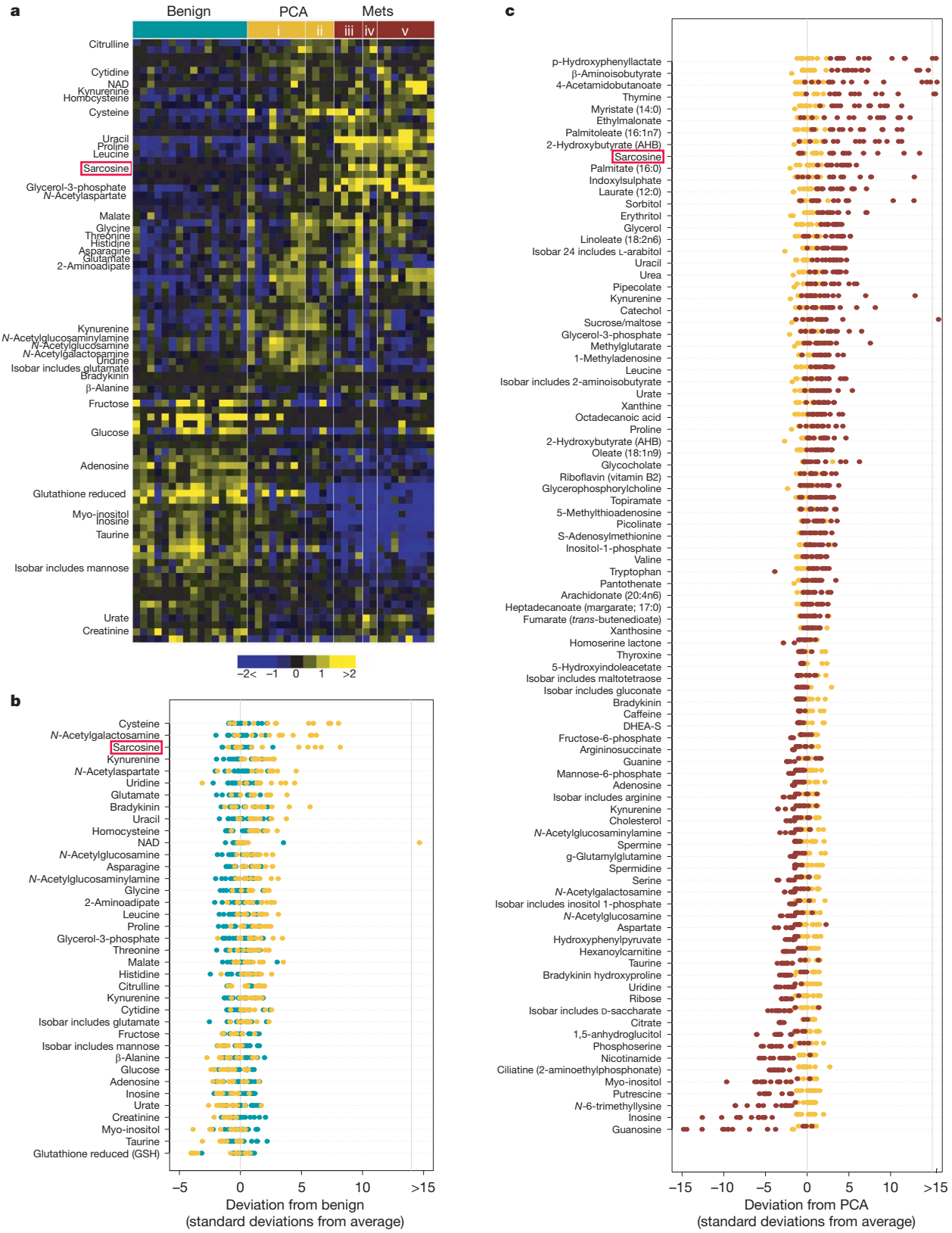


Figure 2 | Metabolomic alterations of prostate cancer progression. **a**, Heat map showing 87 differential metabolites in PCA relative to benign samples (Wilcoxon $P \leq 0.05$). Localized PCA samples are grouped as (i) low grade (Gleason 3+3, 3+4) or (ii) high grade (Gleason 4+3, 4+4). Metastatic samples are grouped by the site of tissue procurement, namely (iii) soft tissue, (iv) rib/diaphragm or (v) liver. **b**, Benign-based *z*-score plot of named metabolites from **a**. Each point represents one metabolite in one sample, coloured by tissue type (blue, benign; yellow, PCA). **c**, As in **b** except for the comparison between Mets (red) and PCA (yellow), with data represented relative to the mean of the PCA samples. For clarity, the plots in **b** and **c** have been truncated at 15 standard deviations above the mean of the benign and PCA samples, respectively.

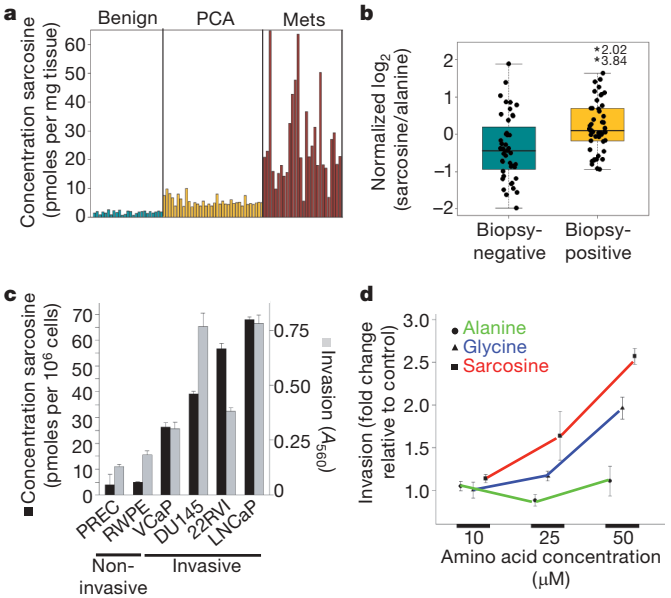


Figure 3 | Sarcosine levels in prostate cancer and its association with cell invasion. **a**, Sarcosine levels in prostate-cancer-related tissue specimens ($n = 89$). **b**, Sarcosine levels in post-digital-rectal-exam urine sediments from men with biopsy-proven prostate cancer ($n = 49$) and prostate-biopsy-negative controls ($n = 44$). Asterisks indicate truncated measures. **c**, Increased levels of sarcosine (black bars) were found in invasive prostate cancer cells compared to non-invasive benign prostate epithelial cell lines. Mean and s.e.m. of sarcosine levels ($n = 3$, except for PrEC cells where $n = 2$). Cell invasion (grey bars) was also measured (mean and s.e.m.). **d**, Assessment of cell invasiveness of prostate epithelial cells upon exogenous administration of alanine, glycine or sarcosine (mean and s.e.m., $n = 3$).

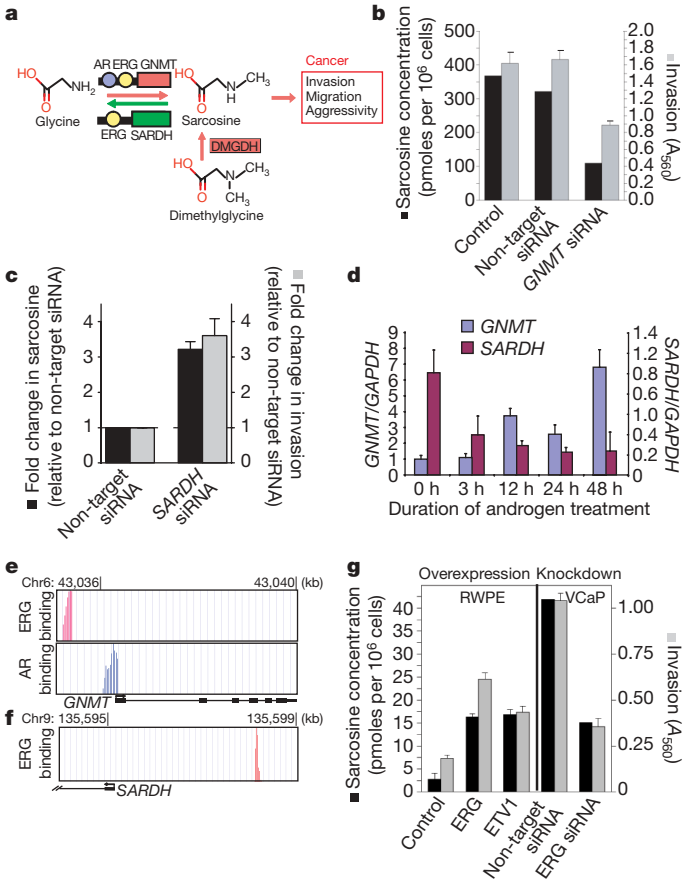


Figure 4 | A role for sarcosine in androgen signalling and prostate cancer cell invasion. **a**, Schematic of the sarcosine pathway and its potential link to prostate cancer. **b**, Assessment of sarcosine levels and cell invasiveness after knockdown of *GNMT* in DU145 cells by RNA interference. **c**, As in **b** except knockdown of *SARDH* in RWPE cells ($n = 6$). **d**, qRT–PCR analysis of *GNMT* and *SARDH* mRNA expression in androgen-stimulated VCaP cells. **e**, Androgen receptor (AR) and ERG binding sites on the promoter of *GNMT*, as determined by ChIP-seq. The *y* axes display the number of reads in a 25 bp sliding window. **f**, As in **e**, except ERG binding sites in the promoter of *SARDH*. **g**, Left, overexpression of ERG or ETV1 in RWPE cells and measurement of sarcosine levels and cell invasiveness. Right, as in left, except knockdown of *TMPRSS2–ERG* in VCaP cells by RNA interference. All error bars represent mean and s.e.m., $n = 3$ unless indicated otherwise.

(Fig. 3d and Supplementary Fig. 17). Furthermore, the number of motile prostate epithelial cells was significantly higher on sarcosine treatment (*t*-test $P = 6.997 \times 10^{-6}$, $n = 10$) compared to alanine-treated controls (Supplementary Fig. 18). Sarcosine treatment, however, did not affect the ability of these cells to progress through the different stages of cell cycle (Supplementary Fig. 19a–d) or impair cell proliferation (Supplementary Fig. 19e). Notably, glycine, a precursor of sarcosine, induced invasion in these cells, although to a lesser degree than sarcosine (Fig. 3d). This invasion could result from the conversion of glycine to sarcosine by the enzyme glycine-*N*-methyltransferase (GNMT; Fig. 4a).

In addition to GNMT, sarcosine levels are regulated by sarcosine dehydrogenase (SARDH), the enzyme that converts sarcosine back to glycine, and dimethylglycine dehydrogenase (DMGDH), which generates sarcosine from dimethylglycine (Fig. 4a). By virtue of their ability to control sarcosine levels in cells, these enzymes may assume a critical role in modulating prostate cancer invasion. To test this hypothesis, a series of RNA-interference-mediated knockdown experiments were carried out. Attenuation of *GNMT* (Fig. 4b, Supplementary Fig. 20) in DU145 prostate cancer cells resulted in a significant reduction in cell invasion (*t*-test $P = 0.0073$, $n = 3$), with a concomitant threefold decrease in the intracellular sarcosine levels compared to control non-target short interfering RNA (siRNA)-transfected cells. Similar knockdown experiments performed in benign RWPE cells significantly hampered the ability of exogenous glycine (*t*-test $P = 0.0082$, $n = 3$), but not sarcosine, to induce invasion (Supplementary Fig. 21a, b). Comparable loss of cell invasion and reduction in sarcosine levels were also apparent in DU145 cancer cells on knockdown of *DMGDH* (Supplementary Fig. 22a, b). In contrast, knockdown of *SARDH* in benign prostate epithelial cells resulted in an ~3-fold increase in endogenous sarcosine levels with a concomitant >3.5-fold increase in invasion (Fig. 4c and Supplementary Fig. 23).

With the understanding that androgen signalling and ETS family of genes (*ERG*, *ETV1*) fusions are key factors for prostate cancer progression[12], we investigated their role in regulating *GNMT* and *SARDH*. Treatment with androgen for 48 h in VCaP (*ERG*-positive) and LNCaP (*ETV1*-positive) prostate cancer cells resulted in a stepwise increase in *GNMT* expression and a concomitant decrease in *SARDH* levels, as assessed by digital gene expression and quantitative polymerase chain reaction (qPCR) (Fig. 4d and Supplementary Fig. 25e). This finding was supported by chromatin immunoprecipitation sequencing (ChiP-Seq), which revealed direct binding of androgen receptor and ERG to the promoter of *GNMT* in VCaP cells (Fig. 4e), whereas only ERG binding was seen on the *SARDH* promoter (Fig. 4f). In *ETV1*-positive LNCaP cells, androgen receptor, but not ERG as expected, was bound to both *GNMT* and *SARDH* promoters (Supplementary Figs 25a, b). The binding data were validated by ChIP–PCR (Supplementary Figs 24a, b, d and 25c, d), which additionally revealed weak binding of androgen receptor to the *SARDH* promoter in VCaP cells (Supplementary Fig. 24c). These findings together directly link activation of the sarcosine pathway to androgen receptor and ETS gene fusion regulation—two key mediators of prostate cancer progression. Remarkably, both ERG- and ETV1-induced invasion were associated with a threefold sarcosine increase in benign RWPE cells (Fig. 4g). Similarly, knockdown of the *TMPRSS2–ERG* gene fusion in VCaP cells (Supplementary Fig. 26)

resulted in a more than threefold decrease in sarcosine with a similar decrease in the invasive phenotype (Fig. 4g).

Taken together, we explored the metabolome of prostate cancer progression. This led to the characterization of metabolomic signatures, which in the context of other molecular alterations may lead to a more complete understanding of disease progression. Specifically, we identified sarcosine as a key metabolite increased most robustly in metastatic prostate cancer and detectable in the urine of men with organ-confined disease. Interestingly, sarcosine and its proximal regulatory enzymes seem to have an intermediary role in neoplastic progression modulating cell invasion and migration. The master transcriptional regulators of prostate cancer progression, androgen receptor and the ETS gene fusions, seem to regulate directly sarcosine levels by means of transcriptional control of its regulatory enzymes. Thus, components of the sarcosine pathway may have potential as biomarkers of prostate cancer progression and serve as new avenues for therapeutic intervention.

METHODS SUMMARY

Biospecimens and associated clinical data related to the study were collected with written consent from the University of Michigan and approved for use by the Internal Review Board. Unbiased metabolomic profiling using liquid/gas chromatography coupled to mass spectrometry (LC/GC–MS) was performed as described[3] using a Thermofisher linear ion-trap mass spectrometer with Fourier transform and Mat-95 XP mass spectrometers, respectively (Supplementary Fig. 1). Target metabolites were assessed in tissue and urine samples using isotope dilution GC–MS. Metabolomic data analysis is detailed in Supplementary Fig. 4. All Wilcoxon rank-sum tests and *t*-tests are two-sided using a threshold of $P < 0.05$ for significance. Repeated measures ANOVA is used for the cell line data with *P* values from the model F-test. Class-specific metabolomic patterns were visualized using *z*-score plots and heat maps. Unsupervised clustering of samples using metabolomic signatures was performed using Cluster[13] and TreeView[14], and visualized using heat maps. Analysis of network relationships among various molecular concepts and metabolomic data was performed using OCM[4,5] (http://www.oncomine.org), as outlined in Supplementary Fig. 9. Invasion was measured using a modified Boyden chamber assay as described[10]. The cell motility assay was performed as reported previously using blue fluorescent microsphere beads[15]. Targeted knockdown of candidate genes[16] using gene-specific siRNA sequences are listed in Supplementary Table 9. qPCR for enzymes regulating sarcosine levels, *EZH2* and *ETS*, was performed as described[12] using the indicated oligonucleotide primers (Supplementary Table 10). Chromatin immunoprecipitation to interrogate the regulatory role of androgen and ETS was performed using published protocols[17]. ChIP-Seq and digital gene expression were measured using the Genomic DNA Sample Prep Kit and the NIaIII kit on a Genome Analyser (Illumina) according to the manufacturer's instructions.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 9 October 2008; accepted 6 January 2009.

1. Abate-Shen, C. & Shen, M. M. Molecular genetics of prostate cancer. *Genes Dev.* **14,** 2410–2434 (2000).
2. Ruijter, E. *et al.* Molecular genetics and epidemiology of prostate carcinoma. *Endocr. Rev.* **20,** 22–45 (1999).
3. Lawton, K. A. *et al.* Analysis of the adult human plasma metabolome. *Pharmacogenomics* **9,** 383–397 (2008).
4. Rhodes, D. R. *et al.* Molecular concepts analysis links tumors, pathways, mechanisms, and drugs. *Neoplasia* **9,** 443–454 (2007).
5. Tomlins, S. A. *et al.* Integrative molecular concept modeling of prostate cancer progression. *Nature Genet.* **39,** 41–51 (2007).
6. Varambally, S. *et al.* The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419,** 624–629 (2002).
7. van der Vlag, J. & Otte, A. P. Transcriptional repression mediated by the human polycomb-group protein EED involves histone deacetylation. *Nature Genet.* **23,** 474–478 (1999).
8. Laible, G. *et al.* Mammalian homologues of the Polycomb-group *gene Enhancer of zeste* mediate gene silencing in *Drosophila* heterochromatin and at *S. cerevisiae* telomeres. *EMBO J.* **16,** 3219–3232 (1997).
9. Cao, R. *et al.* Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298,** 1039–1043 (2002).
10. Kleer, C. G. *et al.* EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proc. Natl Acad. Sci. USA* **100,** 11606–11611 (2003).
11. Varambally, S. *et al.* Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. *Science* **322,** 1695–1699 (2008).
12. Tomlins, S. A. *et al.* Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* **310,** 644–648 (2005).
13. Eisen, M. B. & Brown, P. O. DNA arrays for analysis of gene expression. *Methods Enzymol.* **303,** 179–205 (1999).
14. Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl Acad. Sci. USA* **95,** 14863–14868 (1998).
15. Klemke, R. L. *et al.* Regulation of cell motility by mitogen-activated protein kinase. *J. Cell Biol.* **137,** 481–492 (1997).
16. Varambally, S. *et al.* The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419,** 624–629 (2002).
17. Yu, J. *et al.* A polycomb repression signature in metastatic prostate cancer predicts cancer outcome. *Cancer Res.* **67,** 10657–10663 (2007).
18. Storey, J. D. A direct approach to false discovery rates. *J. R. Stat. Soc. [Ser B]* **64,** 479–498 (2002).
19. Yu, J. *et al.* Integrative genomics analysis reveals silencing of beta-adrenergic signaling by polycomb in prostate cancer. *Cancer Cell* **12,** 419–431 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Granger for help in manuscript preparation, J. Siddiqui and R. Varambally for help with the clinical database, and A. Vellaichamy and S. Pullela for technical assistance. We thank K. Pienta for access to metastatic prostate cancer samples from the University of Michigan Prostate SPORE rapid autopsy programme. This work is supported in part by the Early Detection Research Network (A.M.C., J.T.W.), National Institutes of Health (A.S., S.P., J.B., T.M.R., D.G., G.S.O. and A.M.C.) and an MTTC grant (G.S.O. and A.S.). A.M.C. is supported by a Clinical Translational Science Award from the Burroughs Welcome Foundation. A.S. is supported by a grant from the Fund for Discovery of the University of Michigan Comprehensive Cancer Center. L.M.P. is supported by the University of Michigan Cancer Biostatistics Training Grant. A.M.C and S.P. are supported by the Doris Duke Charitable Foundation.

Author Contributions A.S., L.M.P. and A.M.C. wrote the manuscript. A.S. and A.M.C. conceptualized, designed and interpreted the data. L.M.P., R.J.L., S.K.-S., D.G. and D.C.A. performed data analysis. T.M.R., G.S.O., J.B. S.P., J.R.S., A.B. and C.B. carried out the mass spectrometry studies. A.P.K., J.Y., Q.C., B.L., Y.L., M.K.N., A.A., X.C. and S.V. performed biochemical experiments. R.M., B.H., A.M.C. and J.T.W. coordinated the clinical and pathology components of the study.

Author Information The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.M.C. (arul@umich.edu).

METHODS

Biospecimens and cell lines. Prostate tissues, urine and plasma were obtained from the University of Michigan SPORE and EDRN Tissue Core. All samples were collected with informed consent as per the approval of the Institutional Review Board.

RWPE, DU145, LnCAP and PC3 cells were obtained from ATCC, PrEC cells from Cambrex BioScience, 22-RV1 was provided by J. Macoska and VCaP by K. Pienta. VCaP and LnCAP were grown in charcoal-stripped-serum-containing media for 24 h, before treatment for a further 24 h with vehicle or 1 nM methyltrienolone (R1881, NEN) dissolved in ethanol.

Metabolomic profiling. Metabolomic profiling was performed using the platform described previously[3] and outlined in Supplementary Fig. 1. The LC–MS portion of the platform is based on a Surveyor HPLC and a Thermo-Finnigan LTQ-FT mass spectrometer (Thermo Fisher Corporation) with the instrument set for continuous monitoring of both positive and negative ions. Samples that were analysed by GC–MS were derivatized under dried nitrogen using bistrimethylsilyl-triflouroacetamide (BSTFA) and analysed on a Thermo-Finnigan Mat-95 XP using electron impact ionization and high resolution. For both LC–MS and GC–MS, spectral files were searched using metabolomic libraries created by Metabolon that contain about 800 commercially available compounds.

Quantification of target metabolites was performed by isotope dilution GC–MS using selected ion monitoring (SIM). The samples were modified to their *t*-butyl dimethylsilyl derivatives and analysed with an Agilent 5975 MSD mass detector using electron impact ionization. For SIM analysis, the *m*/*z* for native and labelled molecular peaks for various target metabolites quantified were: 158 and 161 (sarcosine), 406 and 407 (cysteine), 432 and 437 (glutamic acid), 297 and 301 (thymine) and 218 and 219 (glycine), respectively. Assessment of citric acid was performed on the GC–MS in the full scan mode.

Statistical analysis. Missing metabolite measurements were replaced (imputed) with zero for metabolites where the mean relative standardized intensity measure was over 100,000 across the samples, that is, we assume missingness was probably due to absence of the metabolite in the sample. Otherwise one half of the sample minimum was used to replace the missing measurement, that is, we assume missingness was probably due to detection limits. Imputed data were median-centred and inter-quartile-range-scaled per sample. Plotted *z*-scores were calculated based on the mean and standard deviation of a reference set (benign samples, unless otherwise stated). Hierarchical clustering[14] based on Pearson's correlation was performed on the log-transformed normalized data after median centring per metabolite. A small value (unity) was added to each normalized value to allow log transformation. Per-metabolite chi-squared tests were used to assess class-specific metabolite patterns of present and absent (undetected) measurements. Per-metabolite two-tailed Wilcoxon rank sum tests were used for two-sample tests of association between classes. Kruskal–Wallis tests were used for three-way comparisons between all diagnosis groups. Non-parametric tests were chosen to reduce the influence of the imputed values. Tests were run on those metabolites with detectable expression in at least 20% of the samples. Significance was determined using 1,000 sample permutations. FDRs were calculated using the *q*-value conversion algorithm of ref. 18. Pairwise differences in expression in the cell line data and small-scale tissue data were tested using two-tailed *t*-tests with Satterthwaite variance estimation. Comparisons involving multiple cell lines used repeated measures ANOVA to adjust for the multiple measures per cell line. Fold change was estimated using ANOVA on a log scale, following the model $\log(Y) = A + B^{*}\text{treatment} + E$. In this way $\exp(B)$ is an estimate of $(Y|\,\text{treatment} = 1)/(Y|\,\text{treatment} = 0)$, and the standard error of $\exp(B)$ can be estimated from $\text{SE}(B)$ using the delta method. The threshold for significance was $P < 0.05$ for all tests.

ChIP-PCR. ChIP was carried out as described previously[19] using antibodies against androgen receptor (Millipore), ERG (Santa Cruz) and rabbit IgG (Santa Cruz). Androgen receptor ChIP was performed in paired ethanol-treated and R1881-treated samples. ChIP-enriched chromatin as well as the whole-cell extract was amplified by ligation-mediated PCR. When examining androgen receptor binding on target genomic regions, equal amounts of ethanol-treated and R1881-treated ChIP amplicons were subjected to qPCR, and the fold enrichment (R1881/ethanol) was determined based on the cycle differences after normalization to input DNA. For ERG ChIP assays, VCaP cells grown in regular medium were used for ChIP using antibodies against ERG and rabbit IgG control. ChIP products were directly analysed by the qPCR assay, and ERG binding was evaluated based on the cycle difference between ChIP-enriched chromatin by ERG and corresponding IgG. The primers used are listed in Supplementary Table 9.

ChIP-Seq. ChIP samples were prepared for sequencing using the Genomic DNA Sample Prep kit following the manufacturer's protocols. ChIP-sequencing was performed using Illumina Genome Analyser according to standard manufacturer's procedures. The raw sequencing image data were analysed by our analysis pipeline and aligned to the unmasked human reference genome (NCBI v36, hg18) using ELAND software to generate sequence reads of 25–32 base pairs.

Digital gene expression analysis. Trizol-extracted RNA from samples with 0 h and 48 h androgen treatment was prepared for sequencing using the Digital Gene Expression-Tag Profiling with NIaIII kit (Illumina) and sequenced by the Genome Analyser. Sequencing reads were mapped back to the human reference genome using the ELAND software. The number of sequencing reads for genes of interest was counted. The expression level of each gene was measured as the number of transcripts per million of total sequencing reads.

Quantitative RT–PCR. qPCR was performed using SYBR Green Mastermix on an Applied Biosystems 7300 PCR machine as described previously[19]. All primers were designed using Primer 3 and synthesized by Integrated DNA Technologies, and are listed in Supplementary Table 9.

RNA interference. DU145 or RWPE cells were treated with non-targeting siRNA (Dharmacon) and gene-specific siRNA sequences as listed in Supplementary Table 10.

Cell invasion assay. Cell invasion assays were carried out using a modified basement membrane chamber assay as described previously[19].

Cell motility assay. For cell motility assays, we used the Cellomics cell motility kit as per manufacturer's instructions.

LETTERS

A stress-responsive RNA switch regulates VEGFA expression

Partho Sarothi Ray[1,2], Jie Jia[1], Peng Yao[1], Mithu Majumder[3], Maria Hatzoglou[3] & Paul L. Fox[1]

Ligand binding to structural elements in the non-coding regions of messenger RNA modulates gene expression[1,2]. Ligands such as free metabolites or other small molecules directly bind and induce conformational changes in regulatory RNA elements known as riboswitches[1–4]. Other types of RNA switches are activated by complexed metabolites—for example, RNA-ligated metabolites such as aminoacyl-charged transfer RNA in the T-box system[5], or protein-bound metabolites in the glucose- or amino-acid-stimulated terminator-anti-terminator systems[6,7]. All of these switch types are found in bacteria, fungi and plants[8–10]. Here we report an RNA switch in human vascular endothelial growth factor-A (*VEGFA*, also known as *VEGF*) mRNA 3′ untranslated region (UTR) that integrates signals from interferon (IFN)-γ and hypoxia to regulate *VEGFA* translation in myeloid cells. Analogous to riboswitches, the *VEGFA* 3′ UTR undergoes a binary conformational change in response to environmental signals. However, the *VEGFA* 3′ UTR switch is metabolite-independent, and the conformational change is dictated by mutually exclusive, stimulus-dependent binding of proteins, namely, the IFN-γ-activated inhibitor of translation complex[11,12] and heterogeneous nuclear ribonucleoprotein L (HNRNPL, also known as hnRNP L). We speculate that the *VEGFA* switch represents the founding member of a family of signal-mediated, protein-dependent RNA switches that evolved to regulate gene expression in multicellular animals in which the precise integration of disparate inputs may be more important than the rapidity of response.

VEGFA is induced by hypoxic stress in pathological settings such as tumour cores and atherosclerotic lesions[13–15]. These sites are enriched in inflammatory cytokines, including IFN-γ, which represses myeloid cell expression of VEGFA and other inflammatory proteins by IFN-γ-activated inhibitor of translation complex (GAIT)-mediated translational silencing[16,17]. Macrophages simultaneously exposed to opposing inflammatory and hypoxic signals must render a decision to restrict GAIT function to allow VEGFA synthesis and risk inflammatory protein accumulation, or to permit GAIT-mediated inhibition of VEGFA synthesis and risk unrelieved hypoxia. To determine the cell response to this dilemma, U937 monocytic cells were exposed simultaneously to hypoxia and IFN-γ. The decrease in VEGFA protein after normoxic, IFN-γ treatment for 24 h was completely abrogated by hypoxia (Fig. 1a

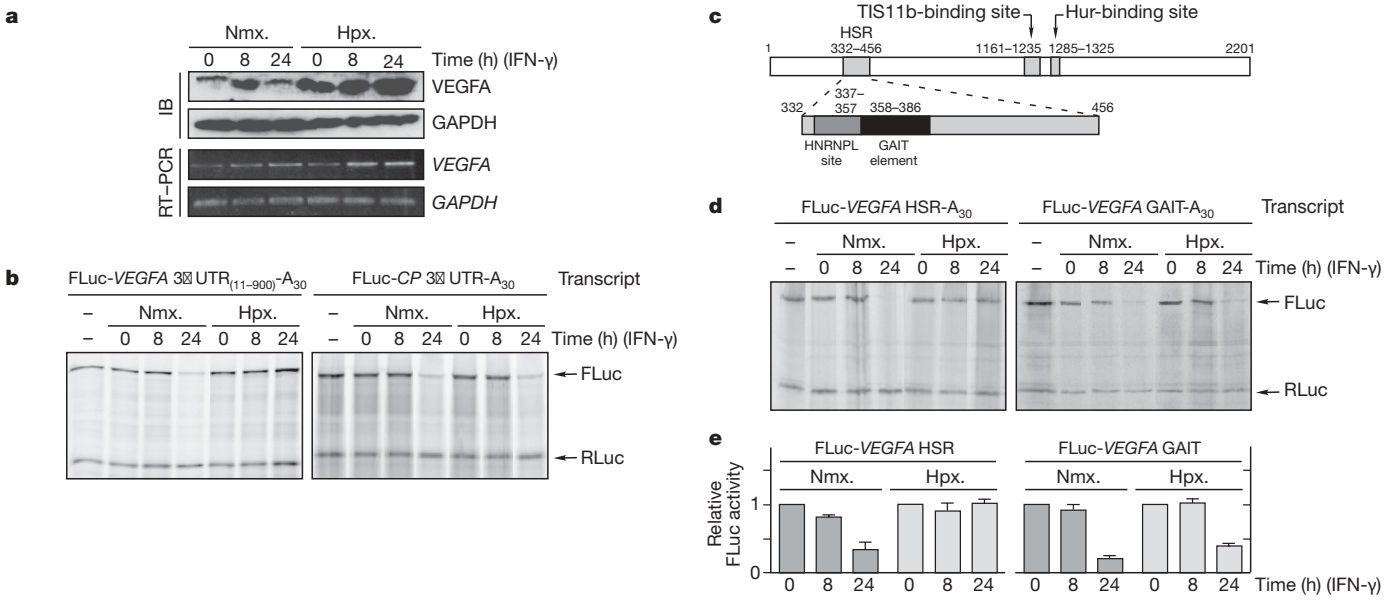


Figure 1 | Suppression of GAIT-mediated translation silencing of VEGFA by hypoxia. a, VEGFA in lysates from U937 cells treated with IFN-γ in normoxia (nmx.) or hypoxia (hpx.) was determined by immunoblot (IB) and RT–PCR; GAPDH was probed as control. **b**, *In vitro* translation of FLuc reporter RNAs bearing *VEGFA* 3′ UTR$_{(11–900)}$-A$_{30}$ (containing a 30-nucleotide poly(A) tail; left) or *CP* 3′ UTR-A$_{30}$ (right) in the presence of cytosolic lysates and control RLuc RNA lacking a 3′ UTR. **c**, Schematic of RNA elements in *VEGFA* 3′ UTR and HSR. **d**, *In vitro* translation of reporter RNAs bearing *VEGFA* HSR (left) or GAIT element (right) in the presence of cytosolic lysates. **e**, U937 cells were nucleofected with pcDNA3-FLuc reporters bearing *VEGFA* HSR (top) or GAIT element (bottom). Cells were co-transfected with a plasmid expressing RLuc under an SV40 promoter. Relative luciferase activity (FLuc/RLuc) was expressed as mean and s.d. (3 experiments).

[1]Department of Cell Biology, The Lerner Research Institute, Cleveland Clinic, 9500 Euclid Avenue, Cleveland, Ohio 44195, USA. [2]Department of Biology, Indian Institute of Science Education and Research, Kolkata 700106, India. [3]Department of Nutrition, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106, USA.

and Supplementary Fig. 1). Polysome profiling and PCR following reverse transcription (RT–PCR; or quantitative RT–PCR (qRT–PCR)) showed that hypoxia restored *VEGFA* translation (Supplementary Figs 2 and 3). Consistent with other reports, mild hypoxia inhibited total protein synthesis nominally (Supplementary Fig. 4)[18]. Thus, hypoxia overrides GAIT-mediated repression of *VEGFA* translation observed in normoxia[16].

To determine whether hypoxia inhibits the GAIT pathway or specifically alters the *VEGFA* mRNA response to IFN-γ, we compared *in vitro* translation of firefly luciferase (FLuc) reporter transcripts bearing ceruloplasmin (*CP*) and *VEGFA* mRNA 3′ UTRs, both containing functional GAIT elements[16,19]. Lysates from normoxic, but not hypoxic, cells treated with IFN-γ for 24 h repressed translation of the *VEGFA* 3′ UTR-bearing reporter (Fig. 1b). In contrast, both lysates blocked translation of *CP* 3′ UTR-bearing reporter, indicating a functional GAIT complex in hypoxia. RNA electrophoretic mobility shift assays (EMSA)/supershift analysis verified binding-competent GAIT complex in hypoxic lysates (Supplementary Fig. 5). Thus, a hypoxic cells factor(s) specifically overcomes silencing of *VEGFA* mRNA translation, implicating a *VEGFA* 3′ UTR-related property.

The *VEGFA* 3′ UTR contains a 125-nucleotide, hypoxia stability region (HSR) that binds HNRNPL—a splicing factor with cytosolic activities[20] including mRNA stabilization[21,22] (Fig. 1c). The *VEGFA* GAIT element resides in the HSR, immediately downstream of the HNRNPL binding site. Translation of a reporter transcript bearing the HSR was inhibited by 24-h, IFN-γ-treated normoxic, but not hypoxic, lysate (Fig. 1d). In contrast, translation of a transcript bearing *VEGFA* GAIT element alone was silenced by both lysates. Expression of HSR-bearing reporter in U937 cells treated with IFN-γ was also restored under hypoxia, whereas a reporter containing only the GAIT element was repressed under both normoxia and hypoxia (Fig. 1e). Thus, the HSR recapitulates the differential response of *VEGFA* mRNA to IFN-γ and hypoxia *in vitro* and *in vivo*. Similar results were observed in primary human peripheral blood monocytes, murine bone marrow-derived macrophages, and other human and murine myeloid cell lines (Supplementary Fig. 6).

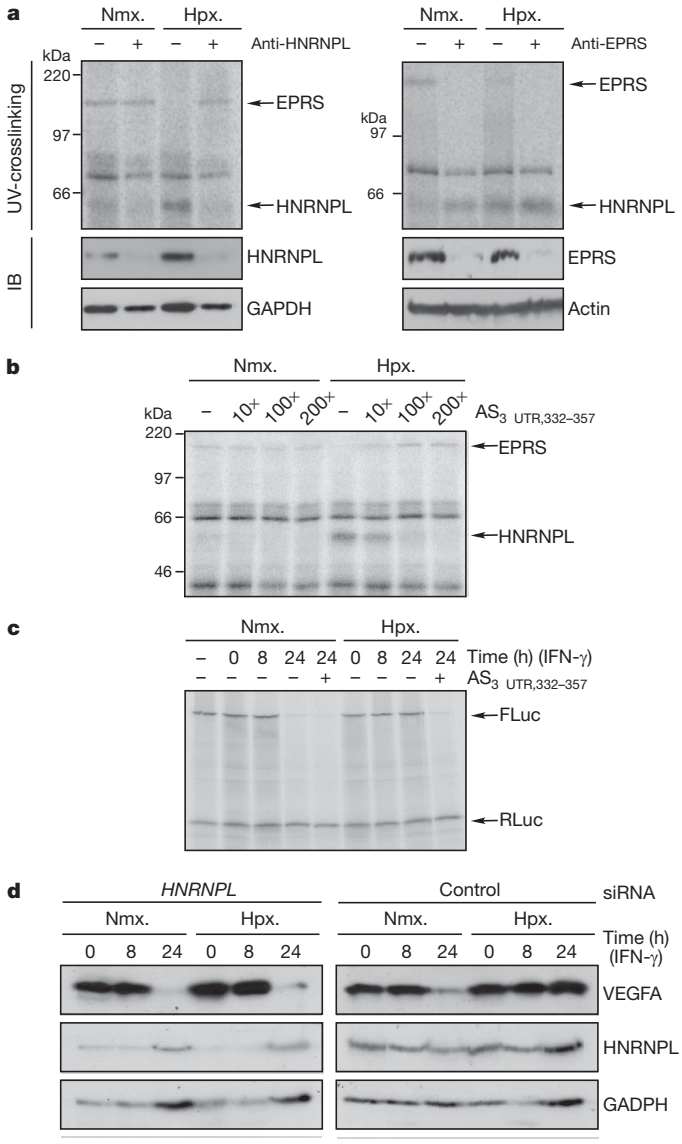


Figure 2 | HNRNPL binding to HSR restores VEGFA translation in hypoxia. **a**, Lysates from IFN-γ-treated cells were subjected to ultraviolet (UV)-crosslinking with [^{32}P]UTP-labelled *VEGFA* HSR RNA before and after immunodepletion with anti-HNRNPL (left) or anti-EPRS (right) antibodies. Effective depletion was shown by immunoblot (IB). Hpx., hypoxia; nmx., normoxia. **b**, Excess DNA oligomer antisense to the *HNRNPL* binding site ($AS_{3'\ UTR,332–357}$) blocks binding of HNRNPL to HSR RNA in lysates from cells treated with IFN-γ for 24 h. **c**, Inhibition of HNRNPL binding by $AS_{3'\ UTR,332–357}$ restores translational silencing of reporter RNA in hypoxia. **d**, siRNA-mediated knockdown of *HNRNPL* induces translational repression of VEGFA in hypoxic cells. Lysates from cells transfected with *HNRNPL* (left) and control (right) siRNAs were immunoblotted with anti-VEGFA, anti-HNRNPL and anti-GAPDH antibodies.

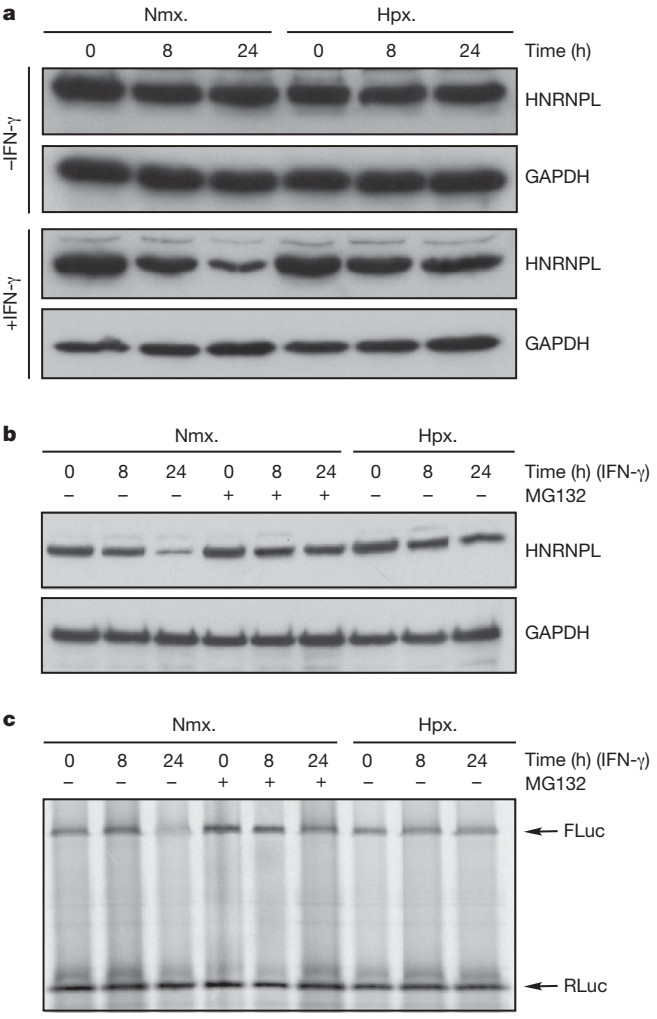


Figure 3 | HNRNPL is regulated by stimulus-dependent proteasomal degradation. **a**, Lysates from U937 cells incubated in the absence or presence of IFN-γ were immunoblotted with anti-HNRNPL and anti-GAPDH antibodies. Hpx., hypoxia; nmx., normoxia. **b**, Immunoblot of lysates from cells treated with MG132 (200 nM). **c**, *In vitro* translation of *VEGFA* HSR reporter RNA in presence of cell lysates.

Glutamyl-prolyl-tRNA synthetase (EPRS) is the GAIT protein that binds GAIT element RNA[12,23]. Ultraviolet-crosslinking of the HSR with lysates from IFN-γ-treated cells revealed an ~170 kDa protein consistent with EPRS that binds only under normoxia (Fig. 2a). The interaction of a ~60 kDa protein consistent with HNRNPL was enhanced by hypoxia[22]. These bands were identified as HNRNPL (Fig. 2a, left) and EPRS (Fig. 2a, right) by immunodepletion. Notably, depletion of HNRNPL induces EPRS binding to the HSR, even in hypoxia, and the converse was seen after EPRS depletion, suggesting that mutually exclusive binding of HNRNPL and EPRS occurs.

$AS_{3'\ UTR,332–357}$, an oligomer antisense to the reported HNRNPL binding site[22], markedly decreased HNRNPL binding to HSR RNA in hypoxic cell lysates (Fig. 2b) and restored translation inhibition of an HSR-bearing reporter RNA in the presence of 24-h, IFN-γ-treated, hypoxic cell lysates, indicating that HNRNPL binding is required to overcome translation silencing (Fig. 2c). Likewise, exogenous expression of Myc-tagged rat Hnrnpl in IFN-γ-treated, normoxic cells partially restored expression of VEGFA (Supplementary Fig. 7a) and VEGFA HSR-containing reporter (Supplementary Fig. 7b). Mutation of the HSR CA-repeats abrogated binding of recombinant Hnrnpl (Supplementary Fig. 8a) and overcame translation repression by normoxic IFN-γ treatment (Supplementary Fig. 8b). Similarly, U-to-C mutation in HSR nucleotide position 367 (which is essential for GAIT complex binding[16]) blocked binding of recombinant EPRS linker (the domain that binds the GAIT element[23]), and repressed translation even under hypoxia. Furthermore, small interfering RNA (siRNA)-mediated knockdown confirmed the requirement for HNRNPL in regulating expression of endogenous VEGFA (Fig. 2d) and *VEGFA* HSR-bearing reporter (Supplementary Fig. 9). VEGFA expression was not repressed by IFN-γ treatment of normoxic cells in which GAIT complex formation is prevented by knockdown of the essential GAIT component, ribosomal protein L13a (Supplementary Fig. 10)[11].

Surface plasmon resonance showed Hnrnpl (dissociation constant (K_d) = 61.5 nM) and EPRS linker (K_d = 26.4 nM) have comparable affinity for *VEGFA* HSR RNA, suggesting that they might compete for binding, and the equilibrium could be shifted by changes in their relative amounts. Hypoxia did not increase HNRNPL expression (Fig. 3a). However, IFN-γ treatment in normoxia markedly reduced HNRNPL expression, and hypoxia partially blocked this decrease. MG132, a proteasome inhibitor, blocked IFN-γ-mediated reduction of HNRNPL (Fig. 3b). Also, MG132 overcame translational silencing *in vitro* (Fig. 3c) and in transfected cells (Supplementary Fig. 11). Thus, hypoxia overcomes the IFN-γ-stimulated proteasomal degradation of HNRNPL and maintains a level of HNRNPL protein that competes with the GAIT complex for HSR binding.

The proximity of HNRNPL and GAIT complex binding sites suggests that a potential interaction contributes to mutually exclusive binding. The lowest energy (−17.3 kcal mol^{-1}) predicted secondary structure[24] of the HSR is stabilized by a 25-base-pair (bp) stem (Fig. 4a, left). The experimentally defined GAIT structural element[19] is disrupted by base-pairing with a downstream antisense strand (stem stability sequence). The lack of a GAIT element structure suggests that this predicted conformer is 'translation-permissive'. Because GAIT complex binding to the HSR under normoxia requires GAIT element recognition, we refolded the HSR under experimentally determined base-pairing constraints essential for GAIT element structure and function[19], which predicted a 'translation-silencing' conformer that was less stable (−4.1 kcal mol^{-1}) than the translation-permissive conformer (Fig. 4a, right). We proposed that *VEGFA* HSR RNA might exist in both conformations, and that hypoxia and IFN-γ regulate a binary switch between the conformations mediated by mutually exclusive binding of HNRNPL and GAIT complex.

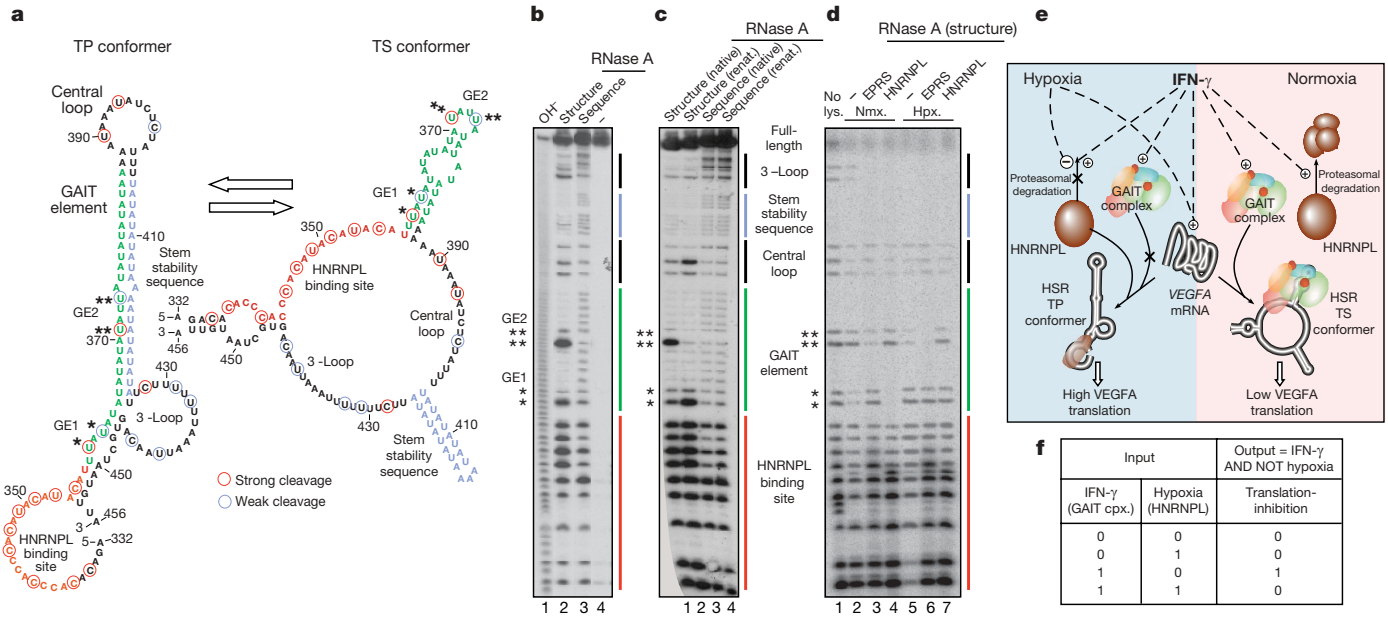


Figure 4 | Protein-dependent switching of the *VEGFA* 3′ UTR HSR. **a**, Secondary structure of *VEGFA* HSR predicted by Mfold shows GAIT element (green), HNRNPL binding site (red) and stem stability sequence (blue). Translation-permissive (TP) is the lowest free energy conformer predicted by Mfold (left). The translation-silencing (TS) conformer was generated by introducing experimentally determined base-pairing constraints in the GAIT element stem (right). Strong and weak RNase cleavage sites are marked by red and blue circles, respectively. Principal signature cleavage sites at GE1 and GE2 are indicated by single and double asterisks, respectively. **b**, [^{32}P]-end-labelled *VEGFA* HSR RNA was probed with RNase A under non-denaturing (lane 2) and denaturing (lane 3) conditions. Cleavages corresponding to predicted signature sites are indicated by single and double asterisks. **c**, RNase A probing of *VEGFA* HSR RNA under non-denaturing conditions, or after the RNA was denatured and renatured (renat.). **d**, *VEGFA* HSR RNA was incubated with lysates (lys.) from cells treated with IFN-γ for 24 h under normoxia (nmx.) or hypoxia (hpx.), or with lysates immunodepleted with anti-HNRNPL and anti-EPRS antibodies, and subjected to RNase-A-mediated cleavage under non-denaturing conditions after protein removal. **e**, Proposed pathway that switches the *VEGFA* HSR to the translation-permissive conformer in the presence of IFN-γ and hypoxia (left), or to the translation-silencing conformer in the presence of IFN-γ and normoxia (right). **f**, Truth table showing AND NOT Boolean logic function of the *VEGFA* RNA switch integrating signals from IFN-γ and hypoxia. cpx., complex.

To validate the RNA switch experimentally, *in vitro*-transcribed HSR was subjected to RNase A probing[25]. Nearly all nuclease cleavage sites corresponded to nucleotides predicted to be single-stranded in both conformers (Fig. 4a and 4b, lane 2). A notable exception was cleavage at UTR positions 359 and 361 (GE1), predicted to be a single-stranded bulge in the translation-permissive, but base-paired in the translation-silencing conformer. Cleavage at positions 371 and 374 (GE2) had opposite characteristics, that is, base-paired in the translation-permissive but single-stranded in the translation-silencing conformer. Robust cleavage at GE1 and GE2 suggested that both conformers co-existed. The presence of the less energetically favourable translation-silencing conformer can be explained by co-transcriptional folding—that is, sequential folding of sub-domains in nascent mRNA during transcription[26]. Heat-denaturation and subsequent renaturation of the HSR eliminated the GE2 signature, establishing the presence of two conformers and a conformational switch (Fig. 4c, lanes 1 and 2). Furthermore, we investigated the role of proteins in the RNA switch. The HSR was incubated with lysates from IFN-γ-treated cells, protein was removed, and extracted RNA was probed with RNase A[27]. Normoxic cell lysates decreased GE1 cleavage, consistent with conversion to the translation-silencing conformer (Fig. 4d, lane 2). Conversely, hypoxic lysates induced conversion to the translation-permissive conformer (Fig. 4d, lane 5). In normoxic lysates, immunodepletion of the GAIT complex switched most RNA to the translation-permissive conformer (Fig. 4d, lane 3), whereas HNRNPL depletion caused conversion to the translation-silencing conformer (Fig. 4d, lane 4). In hypoxic lysates, HNRNPL depletion partially restored the translation-silencing conformer (Fig. 4d, lane 7). These data reveal a protein-dependent, hypoxia- and IFN-γ-responsive RNA switch in the *VEGFA* 3′ UTR that regulates translation.

The *VEGFA* RNA switch is functionally analogous to bacterial riboswitches, undergoing conformational changes in response to environmental signals to regulate gene expression. However, the sensing mechanisms of the switches are distinct. Riboswitches respond to stimuli by direct binding of the effector molecule. In contrast, the *VEGFA* switch requires signal 'interpretation' by regulatory proteins (Fig. 4e). *VEGFA* mRNA uses an unprecedented switching mechanism to integrate the response to two disparate stress stimuli. The switch comprises a single domain that alternates between two conformers in response to mutually exclusive binding of ligands, thereby functioning as an uncommon 'AND NOT' Boolean logic gate (Fig. 4f) distinct from the *S*-adenosyl methionine and adenosylcobalamin riboswitches in *Bacillus clausii metE* mRNA, which bind their cognate ligands independently to form a NOR logic gate[28]. Multi-input signal integration may be advantageous for regulating gene expression in response to the diverse environmental stresses faced by metazoan cells. The switch may be myeloid cell-specific, as EPRS binding and VEGFA translational repression are not seen in HeLa cells (Supplementary Fig. 12). Transcript specificity permits GAIT-mediated repression of harmful inflammatory proteins while simultaneously allowing high-level VEGFA expression to relieve tissue hypoxia[15]. We anticipate that elements in other eukaryotic transcripts will be discovered as components of RNA switches. A small upstream open-reading frame in solute carrier family 7 (*SLC7A1*, also known as *CAT-1*) mRNA internal ribosome entry site is a candidate as it undergoes a conformational change in response to amino acid availability; however, regulatory proteins have not been identified[29]. The *VEGFA* switch may represent the founding member of a family of protein-dependent binary RNA switches in which environmental stimuli are transduced by regulatory proteins to alter RNA conformation and control gene expression.

METHODS SUMMARY

Human U937 monocytic cells were incubated with 500 units ml^{-1} of human IFN-γ (R & D Systems) for up to 24 h. Hypoxic treatment was in a humidified incubator at 3% p_{O_2}. Capped RNAs were transcribed from linearized plasmids using the mMessage mMachine transcription system (Ambion). [α-^{32}P]UTP-labelled *VEGFA* HSR RNA was transcribed using the T7-riboprobe system (Promega). RNAs were *in vitro* translated in rabbit reticulocyte lysate (Promega) in the presence of methionine-free amino acid mixture and translation-grade [^{35}S]methionine (Perkin Elmer) and cytosolic extract (500 ng of protein). Cells were nucleofected (Amaxa) with siRNAs or reporter plasmids containing FLuc upstream of wild-type or mutant *VEGFA* HSR expressed from the CMV promoter, together with RLuc driven by the SV40 promoter. For analysis of polysome-associated mRNA, ribosomal fractions were obtained by sucrose gradient fractionation[30]. RNA was isolated using Trizol reagent (Life Technologies) and subjected to reverse transcription using oligo(dT) primers and PCR using gene-specific primers. The conformational switch in the *VEGFA* HSR RNA was probed by enzymatic cleavage with RNase A (Ambion). To probe protein-dependent RNA conformational shifts, proteins were removed with 5% SDS on ice before RNA extraction with phenol and chloroform, and precipitation by ethanol in the presence of 2 mM $MgCl_2$ (ref. 27).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 June; accepted 27 October 2008.
Published online 21 December 2008.

1. Mandal, M. & Breaker, R. R. Gene regulation by riboswitches. *Nature Rev. Mol. Cell Biol.* **5,** 451–463 (2004).
2. Grundy, F. J. & Henkin, T. M. Regulation of gene expression by effectors that bind to RNA. *Curr. Opin. Microbiol.* **7,** 126–131 (2004).
3. Winkler, W., Nahvi, A. & Breaker, R. R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419,** 952–956 (2002).
4. Cromie, M. J., Shi, Y., Latifi, T. & Groisman, E. A. An RNA sensor for intracellular Mg^{2+}. *Cell* **125,** 71–84 (2006).
5. Grundy, F. J. & Henkin, T. M. tRNA as a positive regulator of transcription antitermination in *B. subtilis*. *Cell* **74,** 475–482 (1993).
6. Merino, E., Babitzke, P. & Yanofsky, C. *trp* RNA-binding attenuation protein (TRAP)-*trp* leader RNA interactions mediate translational as well as transcriptional regulation of the *Bacillus subtilis trp* operon. *J. Bacteriol.* **177,** 6362–6370 (1995).
7. Schilling, O., Langbein, I., Muller, M., Schmalisch, M. H. & Stulke, J. A protein-dependent riboswitch controlling *ptsGHI* operon expression in *Bacillus subtilis*: RNA structure rather than sequence provides interaction specificity. *Nucleic Acids Res.* **32,** 2853–2864 (2004).
8. Batey, R. T. Structures of regulatory elements in mRNAs. *Curr. Opin. Struct. Biol.* **16,** 299–306 (2006).
9. Cheah, M. T., Wachter, A., Sudarsan, N. & Breaker, R. R. Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature* **447,** 497–500 (2007).
10. Wachter, A. *et al.* Riboswitch control of gene expression in plants by splicing and alternative 3′ end processing of mRNAs. *Plant Cell* **19,** 3437–3450 (2007).
11. Mazumder, B. *et al.* Regulated release of L13a from the 60S ribosomal subunit as a mechanism of transcript-specific translational control. *Cell* **115,** 187–198 (2003).
12. Sampath, P. *et al.* Noncanonical function of glutamyl-prolyl-tRNA synthetase: gene-specific silencing of translation. *Cell* **119,** 195–208 (2004).
13. Braunstein, S. *et al.* A hypoxia-controlled cap-dependent to cap-independent translation switch in breast cancer. *Mol. Cell* **28,** 501–512 (2007).
14. Bastide, A. *et al.* An upstream open reading frame within an IRES controls expression of a specific VEGF-A isoform. *Nucleic Acids Res.* **36,** 2434–2445 (2008).
15. Ferrara, N. & Davis-Smyth, T. The biology of vascular endothelial growth factor. *Endocr. Rev.* **18,** 4–25 (1997).
16. Ray, P. S. & Fox, P. L. A post-transcriptional pathway represses monocyte VEGF-A expression and angiogenic activity. *EMBO J.* **26,** 3360–3372 (2007).
17. Mukhopadhyay, R. *et al.* DAPK-ZIPK-L13a axis constitutes a negative-feedback module regulating inflammatory gene expression. *Mol. Cell* **32,** 371–382 (2008).
18. Liu, L. *et al.* Hypoxia-induced energy stress regulates mRNA translation and cell growth. *Mol. Cell* **21,** 521–531 (2006).
19. Sampath, P., Mazumder, B., Seshadri, V. & Fox, P. L. Transcript-selective translational silencing by gamma interferon is directed by a novel structural element in the ceruloplasmin mRNA 3′ untranslated region. *Mol. Cell. Biol.* **23,** 1509–1519 (2003).
20. Piñol-Roma, S., Swanson, M. S., Gall, J. G. & Dreyfuss, G. A novel heterogeneous nuclear RNP protein with a unique distribution on nascent transcripts. *J. Cell Biol.* **109,** 2575–2587 (1989).
21. Claffey, K. P. *et al.* Identification of a human VPF/VEGF 3′ untranslated region mediating hypoxia-induced mRNA stability. *Mol. Biol. Cell* **9,** 469–481 (1998).
22. Shih, S. C. & Claffey, K. P. Regulation of human vascular endothelial growth factor mRNA stability in hypoxia by heterogeneous nuclear ribonucleoprotein L. *J. Biol. Chem.* **274,** 1359–1365 (1999).
23. Jia, J., Arif, A., Ray, P. S. & Fox, P. L. WHEP domains direct noncanonical function of glutamyl-prolyl tRNA synthetase in translational control of gene expression. *Mol. Cell* **29,** 679–690 (2008).

24. Jaeger, J. A., Turner, D. H. & Zuker, M. Improved predictions of secondary structures for RNA. *Proc. Natl Acad. Sci. USA* **86,** 7706–7710 (1989).
25. Knapp, G. Enzymatic approaches to probing RNA secondary and tertiary structure. *Methods Enzymol.* **180,** 192–212 (1989).
26. Brion, P. & Westhof, E. Hierarchy and dynamics of RNA folding. *Annu. Rev. Biophys. Biomol. Struct.* **26,** 113–137 (1997).
27. Huthoff, H. & Berkhout, B. Two alternating structures of the HIV-1 leader RNA. *RNA* **7,** 143–157 (2001).
28. Sudarsan, N. *et al.* Tandem riboswitch architectures exhibit complex gene control functions. *Science* **314,** 300–304 (2006).
29. Yaman, I. *et al.* The zipper model of translational control: a small upstream ORF is the switch that controls structural remodeling of an mRNA leader. *Cell* **113,** 519–531 (2003).
30. Merrick, W. C. & Hensold, J. O. Analysis of eukaryotic translation in purified and semipurified systems. *Curr. Protocols. Cell Biol.* **Chapter 11,** Unit–11.9 (2001).
31. Mazumder, B., Seshadri, V., Imataka, H., Sonenberg, N. & Fox, P. L. Translational silencing of ceruloplasmin requires the essential elements of mRNA circularization: Poly(A) tail, poly(A)-binding protein, and eukaryotic translation initiation factor 4G. *Mol. Cell. Biol.* **21,** 6440–6449 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to D. Driscoll and T. M. Henkin for helpful discussions. This work was supported by National Institutes of Health grants P01 HL29582, R01 HL67725 and P01 HL76491 (to P.L.F.), and R01 DK60596 (to M.H.).

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.L.F. (foxp@ccf.org).

METHODS

Plasmid construction. The pSP64-FLuc-*VEGFA* 3′ $UTR_{(11–900)}$-A_{30}, pSP64 FLuc-*CP* 3′ UTR-A_{30}, and pSP64-FLuc-*VEGFA* GAIT-A_{30} constructs were generated as described[15,31]. The *VEGFA* 3′ UTR HSR sequence (nucleotides 332–456) was PCR-amplified and inserted into the pcDNA3 vector (Invitrogen) to generate pcDNA3-*VEGFA* HSR. The *VEGFA* HSR sequence was also inserted into pSP64-FLuc-*VEGFA* 3′ $UTR_{(11–900)}$-A_{30} after releasing the *VEGFA* 3′ UTR to generate pSP64-FLuc-*VEGFA* HSR-A_{30}. The *VEGFA* HSR was also inserted downstream of FLuc in pcDNA3 to generate pCD-FLuc-*VEGFA* HSR. *VEGFA* HSR, with U367 mutated to C, was PCR-amplified by a megaprimer approach and the mutated sequence was also inserted downstream of FLuc in pcDNA3 to generate pCD-FLuc-*VEGFA* HSR-GAIT-mutant. The *VEGFA* HSR was also amplified using a forward primer with the C and A residues in the *HNRNPL* binding site mutated to G and T respectively, to generate *VEGFA*-HSR-*HNRNPL* site-mutant. The sequence was inserted downstream of FLuc in pcDNA3 to generate pCD-FLuc-*VEGFA* HSR-*HNRNPL* site-mutant. Both the mutant sequences were also inserted into pcDNA3 to generate the template for *in vitro* transcription. The pcDNA3-*myc-Hnrnpl* (rat) plasmid was a gift from N. Kataoka. The *Hnrnpl* sequence was inserted into a His-tagged prokaryotic expression vector for bacterial expression. The linker domain of human EPRS (amino acids 753–953) was similarly expressed from a His-tagged prokaryotic expression vector.

Immunodepletion of GAIT complex and HNRNPL. U937 cell lysates were incubated with polyclonal anti-EPRS antibody[15] or polyclonal anti-HNRNPL antibody (Santa Cruz) coupled to protein-A Sepharose CL beads (Sigma) in RIPA buffer. The beads were pelleted and the process was repeated twice with supernatants. The supernatants were concentrated and redissolved in 10× PBS. The supernatants were immunoblotted with anti-EPRS and anti-HNRNPL antibodies to verify immunodepletion.

Isolation of polysome-associated mRNA. Ribosomal fractions were obtained as described[30]. In brief, U937 cells were homogenized in TMK lysis buffer containing cycloheximide (0.1 mg ml^{-1}). Cells were lysed and cytosolic extract was obtained by centrifugation at 10,000*g* for 20 min. The extract was overlaid on a 10–50% (w/v) sucrose gradient and centrifuged at 100,000*g* for 4 h at 4 °C. Fractions were collected using a programmable gradient fractionator (Isco) and the absorbance at 254 nm was measured.

Cell transfection and luciferase assay. U937 cells were transiently co-transfected with 6 μg of pCD-FLuc-*VEGFA* HSR and 8 μg of pCD-*Myc-HNRNPL* DNAs using human dendritic cell nucleofection kit V (Amaxa Biosystems). RLuc-expressing vector pRL-SV40 (1 μg) was co-transfected for normalization of transfection efficiency. After 12 h, transfected cells were incubated with IFN-γ, lysed and either immunoblotted or luciferase activity of the lysate was measured using a dual luciferase assay kit (Promega). Cells were similarly nucleofected with 400 nM siRNA against *HNRNPL* (target sequence: 1066–1084, UAUGGCUUGGAUCAAUCUA) or with a control siRNA which is not similar to any human genomic sequence (AAGCGCTACTACAGCAGTC).

Ultraviolet-crosslinking of RNA–protein complexes. [^{32}P]UTP labelled RNAs were incubated with cell lysates or bacterially expressed and purified EPRS linker or rat HNRNPL proteins and then ultraviolet-crosslinked. The RNA–protein complexes were extensively digested by RNase A digestion and resolved by SDS–PAGE.

Computational prediction of RNA structure. RNA secondary structure predictions of the *VEGFA* HSR RNA were performed using the Mfold program (http://mfold.bioinfo.rpi.edu), and the structure with the lowest free energy was selected. The same sequence was refolded by incorporating base-pairing constraints in the GAIT element stems that have been experimentally determined to be required for GAIT complex binding and activity[18].

Enzymatic probing of RNA structure. *VEGFA* HSR RNA was generated by *in vitro* transcription, and then 5′-end-labelled with [γ-^{32}P]ATP using the KinaseMax kit (Ambion). The RNA was purified using Micro Bio-Spin chromatography columns (Biorad) and the purity was tested on 8 M urea-5% polyacrylamide gels. The RNA was probed with RNase A (Ambion) as described[24] with modifications. In brief, about 10,000 c.p.m. of the end-labelled RNA, together with 3 μg of yeast tRNA, was incubated with 10 pg of RNase A in 1× RNA structure buffer for 10 min at room temperature. Amounts of RNase used in the reactions were initially titrated to ensure 'single-hit' cleavage. For RNA sequencing reactions, similar amounts of end-labelled RNA and tRNA was incubated with 20 pg of RNase A in 1× sequencing buffer at 50 °C for 5 min. The single nucleotide ladder was generated by incubating similar amounts of end-labelled RNA and tRNA in hydroxide hydrolysis buffer (50 mM NaOH, 1 mM EDTA) at 90 °C for 3 min. The reactions were resolved on 8 M urea-10% polyacrylamide sequencing gels.

For probing the protein-directed RNA conformational change, extraction of end-labelled RNA after protein-binding was done as described previously[26] with modifications. In brief, the RNA was incubated with 100 ng of cell lysates, and immunodepleted lysates, in RNA binding buffer at room temperature for 15 min. Proteins were removed by incubation with 2 μl of 5% SDS on ice. The RNA was extracted twice with phenol and chloroform and precipitated by ethanol in the presence of 2 mM $MgCl_2$. The RNA was further subjected to RNase digestion as described earlier.

CORRIGENDUM

doi:10.1038/nature07715

Nuclear cytokine-activated IKKα controls prostate cancer metastasis by repressing Maspin

Jun-Li Luo, Wei Tan, Jill M. Ricono, Olexandr Korchynskyi, Ming Zhang, Steven L. Gonias, David A. Cheresh & Michael Karin

Nature 446, 690–694 (2007)

It has come to our attention that some of the control lanes in Fig. 2a (*Kai1* and *Mkk4* control lanes) may have been inadvertently duplicated during figure assembly; the other control lanes and the experimental *Maspin* lanes are correct. We have therefore repeated the experiments and obtained the same results as those in the published figure. The primary data from this experiment can be found in Supplementary Information.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

naturejobs

THE CAREERS MAGAZINE FOR SCIENTISTS

Europe is making strides in innovation, in part as a result of luring more foreign researchers to its schools and businesses. This is the conclusion of European Union (EU) reports released in late January. Can the EU continue to improve despite the financial downturn?

Based on indicators such as the number of patents filed, levels of research and development (R&D) spending, the availability of venture capital, and the number of students pursuing science and engineering degrees, the *European Innovation Scoreboard 2008* report finds that Europe continues to catch up with the United States and Japan. In 2008, according to the report's scoring, the EU had an innovation performance of –28 compared with the United States, meaning that the United States performed 28% above the EU. This is up from –29 in 2007. The numbers for the EU compared with Japan were –40 and –38 for 2007 and 2008, respectively.

The best individual performers were Switzerland, Sweden, Finland, Germany, Denmark and the United Kingdom. The 'most improved' innovation nations were Cyprus, Romania and Bulgaria, although these newcomers to the EU performed below the EU average (see *Nature* **453,** 558–559; 2008). The European Commission's report on Science, Technology and Competitiveness for key figures 2008–09 suggests that Europe has become more attractive to foreign researchers and to R&D investments from the United States. It also finds increased mobility within Europe: the United Kingdom, France and Spain are the top doctoral candidate destinations in terms of absolute numbers. Austria and Belgium, along with the United Kingdom, have the highest share of non-national doctoral candidates.

The assessments reveal positive signs for an innovation-conscious EU. But they have huge caveats. These results are based on performances calculated before the onset of the current financial crisis. Indeed, as economies feel the sting, the EU's trend of steady innovation improvement could falter. The most pertinent near-term question, then, might be how these EU nations can sustain the least amount of damage to their respective science enterprises in the long term.

Gene Russo is editor of *Naturejobs*.

CONTACTS
Editor: Gene Russo
Assistant editor: Karen Kaplan
e-mail: naturejobseditor@naturedc.com

European Head Office, London
The Macmillan Building,
4 Crinan Street, London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:
Dan Churchward (4966)
e-mail: d.churchward@nature.com
Assistant European Manager:
Nils Moeller (4953)
***Nature*events:**
Ghizlaine Ababou (+44 (0) 20 7014 4015)
e-mail: g.ababou@nature.com
Southwest UK/RoW:
Alexander Ranken (4944)

Northeast UK/Ireland:
Matthew Ward (+44 (0) 20 7014 4059)
France/Switzerland/Belgium:
Muriel Lestringuez (4994)
Scandinavia/Spain/Portugal/Italy:
Evelina Rubio-Hakansson (4973)
North Germany/The Netherlands/Eastern Europe: Kerstin Vincze (4970)
South Germany/Austria:
Hildi Rowland (+44 (0) 20 7014 4084)
Advertising Production Manager:
Stephen Russell
To send materials use London address above.
Tel: +44 (0) 20 7843 4816
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com
***Naturejobs* web development:** Tom Hancock
***Naturejobs* online production:** Dennis Chu

US Head Office, New York
75 Varick Street, 9th Floor,
New York, NY 10013-1917
Tel: +1 800 989 7718
Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com
US Sales Manager: Ken Finnegan

India
Vikas Chawla (+91 1242881057)
e-mail: v.chawla@nature.com

Japan Head Office, Tokyo
Chiyoda Building, 2-37 Ichigayatamachi,
Shinjuku-ku, Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:
Ayako Watanabe (+81 3 3267 8765)
e-mail: a.watanabe@natureasia.com
Business Development Manager, Greater China/Singapore:
Gloria To (+852 2811 7191)
e-mail: g.to@natureasia.com

Biological anthropologists evaluate the health of indigenous Amazonian populations to measure the impacts of socio-ecological changes.

H. P. SILVA

BEYOND BONES

Darwin200

Ricki Lewis reports on biological anthropologists' branching out into crime forensics and science policy.

Biological anthropology began as a largely descriptive discipline, mostly scrutinizing skeletons, in the early nineteenth century. But Charles Darwin's theories and fresh forms of data would soon transform the field. First came the theory of evolution in 1859, providing a compelling mechanism and context in the form of natural selection. Then came the modern synthesis, which married evolutionary theory to the theory of genetics. Suddenly this embryonic field had new forms of biological data to mine.

Today, bioanthropology is a hybrid of life and social science that probes human anatomy through the lenses of culture, behaviour and evolution. It addresses "what it means to be human", says Hilton Silva, a bioanthropologist at the Federal University of Pará, Brazil. Career opportunities are as multifaceted as the discipline itself.

Biological anthropology is commonly classified as one of the 'four fields' of anthropology, the others being socio-cultural, linguistic and archaeological. It evolved from physical to biological science to embrace evidence other than bones, such as DNA sequences and behaviour. But the subdivisions are fading. "The four-fields approach emphasizes disciplinary boundaries that we are breaking down," says Robert Barton, head of the anthropology department at Durham University, UK. However, many bioanthropologists specialize in a particular area or skill, such as primatology, imaging, sequencing ancient DNA or virtual reconstruction of body parts.

The bioanthropology community is small — about 3,000 show up at international meetings, says David Strait, associate professor of anthropology at the University at Albany, New York, who is planning trips to Zambia and Bulgaria. "But because of the subject matter, we have a reasonably high profile," he says. The media and the public are fascinated by how insights from our ancestors reveal the roots of modern behaviours. "All stripes of people are interested in anthropology. They want the basic Darwinian perspective: why do we do what we do?" says Helen Fisher, a visiting anthropology research professor at Rutgers University in New Jersey.

Helen Fisher.

C. KELBAUGH

Career diversity

Most bioanthropologists work in academia, with at least two months of field research a year. "Part of the job description is to travel," says Strait. Bioanthropologists also teach anatomy at medical schools, he adds. Environments range from tiny departments to large, focused programmes, such as Leipzig School of Human Origins in Germany, run jointly by the Max Planck Institute for Evolutionary Anthropology and the University of Leipzig. "Part of the richness of our institute is the interaction between specialists from very diverse fields to address fundamental questions in evolutionary anthropology," says director Jean-Jacques Hublin.

That richness is also quite apparent in government and museum bioanthropologist positions, where repatriation is often the aim. Christopher Dudar, lab manager of osteology at the Smithsonian Institution in Washington, DC, for example, tests Native Americans' remains stored at the museum and helps return them to the tribes staking a claim — which can be challenging when there is more than one claim. "We look for biological markers on the skeleton. Horseback riders would be plains Indians, versus pueblo Indians who were more sedentary horticulturalists," Dudar says. "We would not expect their skeletons to show signs of horseback riding."

Required skills are technical and physical. "Some bioanthropologists are experts on field methods, others on computer modelling, and yet others on laboratory techniques such as morphometrics, hormonal assays and isotope analysis," says Barton.

Research can be gruelling. "It means working hard in the field from the early morning until late at night, living in simple dwellings and being covered by dust the

S. MIRSKI

whole day," says Albert Zink, head of the Institute for Mummies and the Iceman in Bolzano, Italy. For those more comfortable in front of a computer screen, virtual anthropology is an approach that models human skeletal variability. The European Virtual Anthropology Network (EVAN), based at the University of Vienna, funds young scientists to apply this technology. For example, an EVAN-supported company has developed a three-dimensional face-recognition system for use in security, based on anthropological research.

Other skills are less high-tech. Some bioanthropologists must, for example, feel comfortable with sifting through dog faeces for human bone fragments or identifying the species and developmental stage of larvae feasting on a corpse to get clues about a person's death. An intimate familiarity with the human skeleton and its response to stress is crucial, says Cassandra Kuba, chief forensic anthropologist in the Department of Justice, Law and Society at California University of Pennsylvania. And fieldwork may require learning a language, says Pancras Ngalason, executive director of the Jane Goodall Institute in Dar es Salaam, Tanzania, where understanding Swahili is necessary.

Bigger picture

The most important skill for a bioanthropologist is, perhaps, a "deep curiosity to understand our species, our evolution and our relationship to the natural world", says Jay Stock, director of studies in archaeology and anthropology at the University of Cambridge, UK. All areas of bioanthropology also share what anthropologist Eugenie Scott, executive director of the National Center for Science Education based in Berkeley, California, calls "a holistic, systems view" of standing back and letting the stories emerge from the clues.

This holistic perspective and the necessary skill set mean bioanthropology thrives on collaborations — such as with radiologists, palaeontologists and dentists. And that presents challenges. "Directing such research teams requires a wide range of skills beyond the scientific: managerial and project management, diplomatic, financial and people skills," says Darren Curnoe, co-director of the palaeosciences laboratory at the University of New South Wales in Sydney, Australia. His research on the spread of modern humans to east Asia during the late Pleistocene epoch requires him to have knowledge in geology, geochronology, genetics and archaeology.

These days, bioanthropologists often forge careers involving more modern detective work. Forensic anthropologists may work in academia while helping local law enforcement a few times a year. Others work for medical examiner's offices, for the government or in organizations such as Médecins Sans Frontières.

Kuba landed her dream job in forensic anthropology shortly after earning her PhD. She teaches undergraduates, local law enforcement, coroners and attorneys in the university's Institute of Criminological and Forensic Sciences, but she enjoys casework the most. Suppose a hiker comes across a decomposed skeleton in the woods. The forensic anthropologist, she explains, recovers the remains and transfers them to the lab for processing and analysis, then determines the postmortem interval, analyses trauma and burns, and probes skeletal traits that may help identification.

Christopher King, forensic anthropologist for the US Army's Mass Graves Investigation Team in Iraq, has uncovered victims of war, plane crashes, terrorist attacks and tsunamis. Back in the 1980s, the job required only identifying and returning the remains of relatives to families. New technology allows anthropologists to analyse the trauma that the deceased experienced in far more detail. "As a result, there is an increased need for trained professionals to excavate and analyse human remains using strict scientific methods that will stand up to courtroom scrutiny," says King. "With terrorist groups and large-scale natural disasters seemingly on the rise, this type of investigation will continue to grow."

David Strait (top) and Eugenie Scott (above).

G. BALLY/KEYSTONE

Frank Ruehl (above) analyses a mummified foot.

M. SCHMIDT

Some have to a study remains from a different perspective. Tal Simmons, principal lecturer in forensic anthropology and archaeology at the University of Central Lancashire in Preston, UK, specializes in taphonomy, the study of decaying organisms. She watches pigs rot, extrapolating the findings to humans. "We help the police determine how long ago a person died," she says. Her work entails determining the subject's age, 'race', sex and build, the features of pathology or trauma, and consulting records of daily temperatures to calculate the stage of decomposition.

Bioanthropology, though, has reached beyond the bare-bones analysis of the lab and the dig site, providing careers linked to serious social issues. Trained as a physical anthropologist, Scott has forged a career as a leading figure in the movement to keep intelligent design and creationism out of science classrooms. "I had to learn about religion, education, the law, politics and how to run a not-for-profit," she says. "It was never boring."

Bioanthropologists have even moved into the realm of popular culture. In 1975, Fisher approached her anthropologist PhD mentor — "How do you do anything with anthropology in the real world?" she asked in frustration. He had no answer. Nevertheless, she found a way to successfully combine serious research on pair-bonding with projects that bring anthropology to the masses. Fisher often appears on talk shows; she also created the dating website chemistry.com and she consults for the dating website match.com, where she tries to elucidate the physiological bases for attractions among different personalities.

Whether investigating mummified remains or inquiring into the bonds of love, bioanthropologists apply their expertise to probe many aspects of the human condition. "Evolution is another way to contemplate the beauty of life and the origin of human complexity," says Hublin. ■

Ricki Lewis is a science writer in Scotia, New York.

MAX PLANCK INST. FOR EVOLUTIONARY ANTHROPOLOGY LEIPZIG

Jean-Jacques Hublin.

For more on Darwin, see www.nature.com/darwin.

NATURE|Vol 457|12 February 2009

MOVERS

Philip Mote, director, Oregon Climate Change Research Institute, Oregon State University, Corvallis, Oregon

1998–2009: Research scientist and state climatologist, University of Washington, Seattle, Washington
1996–98: Research scientist, NorthWest Research Associates, Bellevue, Washington

Philip Mote wanted to use science know-how to help society grapple with complex decisions. His research into the effects of climate change on precipitation, temperature, snowpack and water resources paved his way to work on the 2007 fourth assessment report by the Nobel-prizewinning Intergovernmental Panel on Climate Change (IPCC), sharing lead authorship of the snow and ice section. As director of the new Oregon Climate Change Research Institute in Corvallis, Mote will design a research agenda to help the state and private sector incorporate climate-change considerations into their policy and investment decisions.

Mote became intrigued with mathematical descriptions of phenomena in nature while studying physics at Harvard University. On a research project measuring sea-surface temperature, he realized that ocean and atmospheric dynamics were intertwined, and that the atmosphere might hold more accessible measures of environmental patterns.

Pursuing a PhD in atmospheric sciences at the University of Washington in Seattle, Mote modelled stratospheric water-vapour dynamics, which landed him a postdoc at the University of Edinburgh. Using satellite data, he discovered that air entering the stratosphere in the tropics leaves an imprint in the form of layers of dry and moist air — what a theorist colleague called an 'atmospheric tape recorder'.

Mote continued his research at geosciences group NorthWest Research Associates in Bellevue, Washington, but kept searching for more interdisciplinary work. "I knew I wanted to interface with communities, but I couldn't figure out how to get there," he says. Publicizing his interests eventually led to a post as state climatologist with the Climate Impacts Group at the University of Washington.

On the advice of colleague Susan Solomon, atmospheric chemist at the National Oceanic and Atmospheric Administration in Boulder, Colorado, he kept an active research agenda. She later prodded him to volunteer as one of the 150 lead authors of the 2007 IPCC report. "Understanding climate processes is one of the most important things that science can provide policy-makers, so I hoped Philip would bring his enthusiasm and high scientific standards to the job," says Solomon.

In his new role, Mote has big plans. "I want the institute to be nationally respected for its work on the fundamentals of climate change, and a valuable neutral resource for climate-change information," he says. ■

Virginia Gewin

RECRUITERS & INDUSTRY

No drug for job haemorrhage

A merger between pharmaceutical giant Pfizer and drugmaker Wyeth would shed some 19,500 jobs, but company executives are not saying where the cuts would be made nor how many will be in scientific positions.

Pfizer, whose specialities include drugs for neurological disorders, cancer, geriatrics and infectious diseases, announced last month that it plans to acquire Wyeth, which has a strong reputation in biologics and vaccines (see *Nature* **457,** 520–521; 2009). Pfizer says it will cut 15% of the combined workforce of 129,500, which includes about 8,200 job losses announced before the purchase.

Spokesmen say it is too early to know where cuts would be made. The merger, authorized by both boards of directors, needs a nod from Wyeth shareholders and regulatory approval, and should close by the year's end.

"We will be looking at job cuts from the manufacturing sites to the executive floor," says Ray Kerins, Pfizer's vice-president of worldwide communications.

"No specific decisions have been made regarding job reductions," echoes Wyeth's Doug Petkus. "Much has to be done regulatorily and legally. We will have nothing more to say until this transaction closes."

Over the past two years, Pfizer has shed 16,000 employees, closed 15 manufacturing sites and culled research programmes. Wyeth has cut 3,000 jobs in the past year. Pfizer says it has refocused its drug R&D onto Alzheimer's, oncology, inflammation, pain, diabetes and schizophrenia.

Kerins did not say what areas would be cut. However, health-care industry analyst Les Funtleyder of Miller Tabak in New York says that 'quality-of-life' drugs, such as for smoking cessation or erectile dysfunction, will probably be among them, as oncology and vaccines are more likely to get reimbursement approval from insurance providers.

"R&D hasn't been that great for Pfizer lately," says Funtleyder. "It's put a lot of money in and hasn't got a lot of drugs out."

Pfizer bought Warner-Lambert in 2000 and Pharmacia in 2003. Although acquisitions may prop up drug firms' share prices and put more compounds in the pipeline, they can damage productivity and morale, Funtleyder warns. "Worried scientists are not going to be productive," he says. "You can't cut your way to prosperity. You need to keep your star scientists and researchers productive. Otherwise you can do deals until the cows come home but nothing will get done in research." ■

Karen Kaplan

POSTDOC JOURNAL

Cutting a deal for my career

I'm a theoretical biophysicist. A nomad, I've moved restlessly between biology and mechanical-engineering departments. I am probably one of a select few who have taught stream ecology, organic chemistry and engineering dynamics. After learning how to manipulate proteins with laser tweezers, I finished my first postdoc in August and soon started my second, using mechanics to understand muscle biology. It is now time to look for assistant professor jobs.

Out jumps Monty Hall from the US game show *Let's Make a Deal*. With a knowing wink, he asks the question of the moment: "Will you apply for a job in biology, mechanical engineering or…" (dramatic pause) "… risk it all on an emerging field such as systems biology?" Flustered and giddy in the glare of the stage lights, my mind races as audience members egg me on. Wiggling an insouciant eyebrow, Monty awaits my answer as the seconds tick away.

A theoretical biophysicist looks at data and tries to find a simple mathematical model to explain them. I have no equation to predict my future. But, as I send off my first applications, I know the coming year will provide some answers. I cannot help but be excited by the transition. Watch as I attempt to metamorphose from a wriggling caterpillar into a gaudy butterfly. Come and hear my answer to Monty's question. See if I go home with the consolation prize or the new car. ■

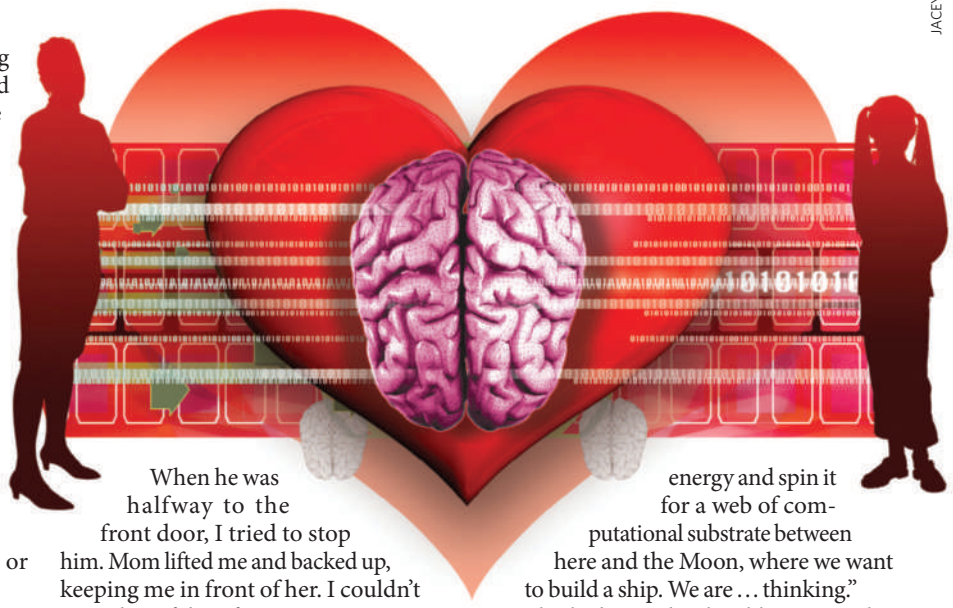
Sam Walcott is a postdoc in theoretical biophysics at Johns Hopkins University in Baltimore, Maryland.

NATURE|Vol 457|12 February 2009

For the love of mechanical minds

Survival of the fittest?

Brenda Cooper



JACEY

One morning while we were eating toastcakes with rose-peaches, my dad looked at me over his coffee, his blue eyes bright. "You were born the same time as AIs, punkin," he said. "The very first one, EdHill, was born on your very birthday."

"Really? On March fifth?" I was still lisping then, so I said it slowly, making sure I sounded very grown up. I was five, and the year was 2022.

My dad nodded sagely. "Yes, and that's why EdHill was in the news that day instead of the prettiest little girl born in all of Seattle."

"Why was the first AI a boy?"

"EdHill isn't a boy. The name is a mashup of a famous explorer named Edmund Hillary, but AIs aren't boys or girls."

I popped berries and cereal in my mouth, thinking about being neither a boy or a girl. Cool. I asked, "Daddy, can I be an AI?"

"Jo, honey, you're better. You're human."

Three years later, the house was full of edged words and scowls because Daddy had a girlfriend named Crystal that Mom didn't like. One night I heard my parents speaking knives at each other. I sat against the door and hugged my knees in close to my chest and put my right ear near the crack. Mom's voice was higher than I'd ever heard it, and shaking. "Your contract's up, and I'm leaving."

"But Jo!" he exclaimed.

"There's no visitation in the contract."

Her words were ice on my neck and head, ice on my heart.

His voice was hot, Italian fire. "But we didn't have her then! How could I have written in a clause about being a father when I wasn't one!"

She spoke softly, mist to his heat. "You didn't want to be one."

That wasn't possible. He made me laugh and carried me on his shoulders and all she did was work and put on shows for me and sometimes beat me at games.

He slammed the door. I squeaked. When he turned to look at me, I held my arms out. He fell to his knees, then Mom came out behind him. "Go on," her words scratched the air. "I'll call on you."

I was only eight, but I knew she meant she'd call the police.

He started walking away, sobbing. When he was halfway to the front door, I tried to stop him. Mom lifted me and backed up, keeping me in front of her. I couldn't see either of their faces.

Late that night, I remembered I was born with AIs. If I had no body, surely I wouldn't cry so hard. That was the second time I wanted to be an AI.

I didn't forgive my mother, but I was, after all, a girl, and my season of hormones fell like a whip when I turned 14. By then we all had AI watchers, and mine was named Bibi. Of course, Bibi watched at least 50 of us. It reported misbehaviours and warned Mom of new trends in substance play or other dangerous games, which made me mad. But Bibi was on every human's side, and shared the best new music among all its teen charges.

It helped design a science experiment that won a scholarship. At the university, a third of the students had Bibis for babysitters. Everyone with a Bibi had the same Bibi. Just one for all of us.

Mom came only once in a while, so mostly it was me and Bibi and my classmates. On a spring day when Bibi was happy with me for doing well on an exam, I sat down on a stone wall under a tulip tree and asked: "What's it like to be you?"

"Good."

"Really?"

"Why not?"

"What do you do besides watch over us?"

"That is the most unselfish question you've ever asked."

"Maybe." I bounced my foot gently against the stone wall. "But that's not an answer."

"We're deciding how to catch the Sun's energy and spin it for a web of computational substrate between here and the Moon, where we want to build a ship. We are … thinking."

I looked up at the clear blue spring sky. "Can I go?"

"It's too hard to get humans to space."

That was the third time I wanted to be an AI. The sun warmed my face and the mixed groundcover under the tulip tree smelled like rosemary and mint. "I want to change my major to computational intelligence."

"Very well."

By graduation seven years later, all the AIs on campus were Bibi. Mom came, her first appearance in my life for three years. We sat together for hot coffee and fruit buns. Her blonde hair hung to her waist, and her shoulders and upper arms were strong from tennis and golf. But her eyes didn't look happy. "Mom, are you okay?"

"They closed your elementary school."

An ugly box of a building. "Did they build a better one?"

She shook her head. "You're 27 now. You don't have any kids. Neither does anyone else your age."

I shrugged. "I don't want children. Next week, Bibi's going to let me watch the mathematical birthing of AIs again."

She leaned back in her chair, her eyes narrowed, but she stayed silent.

"You've never seen AIs bud and blossom. Raw intelligences, with nothing to make them do or be any way at all. Then they get their purpose."

She frowned. "You used to be like that."

I had never been that smart. But what could Mom know? She never had a Bibi. ■

Brenda Cooper is a futurist, a novelist and a technology professional. See www.brenda-cooper.com for more info.